NOAA Technical Report NESDIS 51

THE IMPACT OF NOAA SATELLITE SOUNDINGS ON THE NUMERICAL ANALYSIS AND FORECAST SYSTEM OF THE PEOPLE'S REPUBLIC OF CHINA

Part I - FGGE Data

Part II - December 1987 Data

Co-Principal Investigators:

Dr. Arnold Gruber - NOAA/NESDIS
Prof. Wang Zonghao - SMA/SMC

Co-Investigators:

Mr. Albert Thomasell - NOAA/NESDIS
Dr. T.S. Chen - NOAA/NESDIS
Mr. Irwin Ruff - NOAA/NESDIS
Mr. Mao Jianping - SMA/SMC
Mrs. Huang Ji-hong - SMA/SMC
Mr. Dale Kaiser - SSAI

Washington, D.C.
May 1990

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary

National Oceanic and Atmospheric Administration
John A. Knauss, Under Secretary

National Environmental Satellite, Data, and Information Service
Thomas N. Pyke, Jr. , Assistant Administrator

Contents

Table

Figures

Abstract

The impact of NOAA satellite sounding data on the numerical analysis and forecast system of the People's Republic of China has been investigated. The results are presented here in two parts. Part I describes the result of 1979 FGGE data, and part II, the December 1987 data.

Four experiments were conducted. They are: (1) NOSAT; which uses conventional data only and provides benchmark statistics for the performance of the forecast model, (2) SAT; which includes satellite and conventional data (3) ALLSAT; which uses satellite data only, except for conventional surface data required to define the 1000 mb height reference level, and (4) SATGUESS; which uses the ALLSAT analysis as the first guess to which conventional data are then added.

Seven areas deemed to have adequate conventional data for definition of "truth", including Europe, West China, East China, North America, West Pacific, East Pacific and North Atlantic were chosen to compute the verification statistics. The impact of satellite data was then assessed by computing systematic errors, changes in S1 skill score and root mean square errors for 24 and 48 h forecasts of Sea Level Pressure (SLP) and 500 mb height.

Overall, the Chinese model is found to underforecast the strength of features in the mean SLP fields, i.e., pressures are too low near high pressure systems and too high near low pressure systems. The same type of pattern is also seen in the mean 500 mb height forecast. In general, satellite data have been shown to

reduce systematic error near where the data have been introduced in the vicinity of low pressure systems.

A positive impact on root mean-square errors of up to 11.5% was observed over land for SLP and 8% over ocean for 500 mb height for forecast period of 48 h. However, only a meager increase of 5.5% in S1 skill score was observed at SLP for 48 h forecast over oceans. Some areas show little or no improvement.

Part I - FGGE Data

1. Introduction

In 1987, approval was obtained from the U.S. - PRC Protocol on Atmospheric Science and Technology to conduct a study on the impact of satellite data on the numerical analysis and forecast system of the Peoples Republic of China. The analysis and forecast system used by the PRC does not currently utilize satellite data. Prior impact studies e.g., Tracton et al., 1980, Bengtssen et al., 1982 and Halem et al., 1982 show little or no impact to significant positive impacts on forecasts. An earlier impact study conducted with the Israel Meteorological Service (IMS), (Wolfson et al., 1985, Thomasell et al., 1986) showed that satellite sounding data had a significant impact on 24 and 48 hour forecasts. The IMS model, however, is relatively unsophisticated, particularly in its analysis cycle which used a 12 hour old analysis as the first guess, so that one might expect to see significant impacts, as discussed by Tracton, et al., 1980. The PRC analysis forecast system is much more sophisticated, and conducting an impact study with their model should provide some insight on the dependency of the impact on the analysis forecast system.

The experiments planned are similar to the ones conducted in the IMS study. They are: (a) NOSAT; which uses only conventional data and will provide the forecasts and analyses from which impacts will be measured, (b) SAT; which includes satellite and conventional data, (c) ALLSAT; which uses satellite data only, except for conventional surface data required to define the 1000

mb height references level, and (d) SATGUESS; which uses the ALLSAT analysis as the first guess to which conventional data are then added. It was the SATGUESS experiment which gave the largest impact in the IMS study. Two data sets are planned; one from the FGGE in January 1979 (Part I), which is from a one satellite system and the other is from December 1987 which contains soundings from two satellites, and which also has the advantages of improvements made to the sounding retrieval system, McMillin and Dean (1982).

2. Data

The data used for this study are the FGGE Global II-b level data for the period of January 6 through 30, 1979. Both 00 GMT and 12 GMT of this data period were used to produce analyses and forecasts out to 72 hours. Depending on the type of experiment, either conventional data or satellite soundings or the combination of the two are used. For example, the basic experiment uses only conventional data to produce analyses and forecasts. The forecasts are then verified against the "truth" which is derived from the conventional data analysis. In the case of SAT experiment, satellite soundings in addition to the conventional data are used. However, the satellite soundings are only used over the ocean regions.

In this study, conventional data include surface pressure, and temperature, dew-point depression, wind and height from radiosonde reports at each mandatory level. Sometimes aircraft winds and cloud motion winds are also included.

The satellite soundings comprise temperature and height at

mandatory levels and are only available from one satellite - TIROS N during the experiment period. To obtain adequate satellite data coverage over the oceans, it was necessary to accept data over a broad period, + six hours to - three hours, with respect to map time. The window for other non-conventional data such as aircraft winds and cloud winds was ± three hours. This data set is in FGGE format. The December 1987 data set is in ON 29 format.

Satellite soundings are converted from layer thickness values to heights at mandatory levels by adding the 1000 mb height. The 1000 mb height field for the SAT experiment may be obtained from both surface data and radiosonde data. For the ALISAT experiment only surface data may be used. The accuracy of this conversion has always been a significant statistical problem.

3. Model

Experiments with numerical models in the past show that the impact of satellite soundings depends not only on how the soundings are used by the models but also on the inherent system of the model as well (Thomasell et al., 1986; Mao, 1986). This is because a model first objectively analyzes the initial fields of the meteorological parameters and then produces the forecast by integrating these initial fields with respect to time. Thus, the impact of satellite data on weather prediction clearly depends on the data assimilation technique and the ability of the model to describe the physics of the atmosphere.

The numerical model which has been in operational use by the Beijing Meteorological Center (BMC) was adapted to generate

analyses and forecast fields for this impact study. It is a primitive equation model with a combination of σ -p coordinates operating at five sigma levels defined on a total of 51 x 51 grid points in the Northern Hemisphere with a grid mesh of 381 Km true at 60° N. The σ -p coordinate is a modified σ - coordinate using the p-coordinate in the upper 600 mb and σ - coordinate in the lower 600 mb (lower troposphere). The model uses a 12-minute time step. The so called Cressman's successive correction scheme (1959) was adopted for objective analysis. There are ten analysis levels; surface and all mandatory levels up to 100 mb. Pressure, temperature, dew point depression, wind and height fields are all analyzed at these levels. The boundary is treated as a sponge layer allowing a zero gradient flow. The model includes some physical parameterization such as surface friction, orography, horizontal diffusion and diabatic heating.

4. Experiments

Four experiments were designed. They are similar to the experiments jointly conducted by the Israel Meteorological Service and NESDIS (Thomasell et al., 1986). The experiments consist of producing analyses and forecasts up to 72 hours with the Beijing Meteorological Center (BMC) analysis and prediction models, using the Global II-b level data. The data set for January 1979 was in the FGGE period and comprised conventional data and satellite soundings from one polar orbiting satellite.

The first experiment, called NOSAT, uses conventional data only and provides benchmark verification statistics for the

performance of the BMC numerical forecast model. Conventional data consist of standard surface and ship reports, aircraft winds, and radiosondes. The NOSAT analyses are used throughout all the experiments for verifying forecasts. Over land in dense data regions they are a good representation of the truth. In data-sparse regions, especially over oceans, Tibetan plateau and Greenland, they will not be a good measure of truth and care must be taken in their use. The second experiment, called SAT, uses both satellite and conventional data with equal weighting. The satellite soundings, however, are introduced only over the ocean regions. A comparison of the NOSAT and SAT experimental results gives a measure of impact. A third experiment, denoted ALLSAT, was designed to test the performance of an all-satellite-data system. In this experiment only satellite data and conventional surface data are used. The surface data are required to construct the 1000 mb height that is used as a reference level for the satellite soundings. The final experiment, denoted SATGUESS, was an attempt to optimize the combination of satellite and conventional data. Here, satellite data, actually the ALLSAT analysis, are used to define a first guess for the analysis which is then modified by conventional data in the same manner as the NOSAT experiment. This procedure yields analyses resembling those of ALLSAT in normally data-void regions and resembling NOSAT in areas of dense or adequate radiosonde data. Overall, all experiments begin with climatological data as the initial field. After 24 hours, the initial guess fields are corrected by the forecast values. The

weight of climatology and forecast values are assigned depending on the location of each grid point. For example, a full weight is assigned to climatology at grid points north of 30°N, whereas the forecast fields are given a heavier weight to grid points south of 30°N.

A data set combining TIROS-N satellite soundings and conventional data for the period of 6-30 January 1979 was used, and the BMC primitive equation model was adapted to make the forecast. Evaluation of the experimental results was accomplished objectively through the computation of verification statistics. Figure 1 shows the forecast domain of the BMC model and seven verification areas. There are four land areas, Europe, West China, East China and North America, and three ocean areas, West Pacific, East Pacific and North Atlantic. Verification is then done by comparing a forecast field with a best estimate of the true field. In regions where the analysis is determined by dense, high quality radiosonde data, the NOSAT analysis provides a very good representation of the truth. The primary verification statistics are the root-mean square difference between a forecast field and its verifying NOSAT analysis and the corresponding S1 skill score (Teweles and Wobus, 1954) for each of the seven verification areas. Sea level pressure and 500 mb height verification statistics are given separately for the land, ocean and also for the combination of land and ocean areas.

Another interesting experiment is to examine the impact of satellite data on the systematic error of the BMC model. This can

be done by defining the mean or systematic forecast error E_x at a grid point p by

$$E_x\ (P) \ = \frac{1}{N} \sum_{N=1}^{N} [\ X_f\ (p,\ N) - X_a\ (P,\ N)\ ], \qquad (1)$$

where X_a is the verifying analysis value (in this case NOSAT), X_f is the forecast value and N is the sample size or number of maps over which the average is calculated. Equation (1) may be rewritten as

$$E_x\ (P) = \overline{X_f}\ (P) - \overline{X_a}\ (P), \qquad (2)$$

and a field of systematic errors may be computed as the difference between the average forecast field and the average analysis

A determination of the statistical significance of difference between NOSAT, SAT, ALLSAT and SATGUESS systematic forecast errors was made by applying Student's t test. A review of this test and an explanation of how it may be properly applied to a study of this nature can be found in Wolfson et al. (1985).

5. Results

To measure the performance of the BMC model and impact of satellite data, two objective schemes are adapted to verify the forecast against the observation. The first quantitative measure is S1 skill score which was designed by Teweles and Wobus (1954). In summary, the S1 skill score unit is in percentage and can be larger than 100 for a very poor forecast. The lower the score is, the more skillful the forecast. According to an investigation by Halem et al., (1982), the scores of useful forecasts for sea level pressure and 500 mb height are less than 80 and 60, respectively. The second measure is the root-mean square error. These two

quantities are used to verify forecasts against analyses over the chosen regions and are the basis of our ensuing discussions. Bear in mind that they have basic differences in nature so that they may not give the same sign for the impact of satellite data on a forecast. This is illustrated in Table 1, which shows typical verification statistics for 48 hour forecasts over land.

5.1 S1 Skill Scores

Fig. 2 shows the temporal average of 24 hour forecast S1 skill score at sea level pressure (SLP). For each verification area, it is evident that the SAT experiment shows either negative or negligibly small positive impact over NOSAT experiment. The worst result, however, is in the ALLSAT experiment. This impact relationship also holds true for results at both 48 and 72 hours, particularly at 72 hours where the skill score of each experiment converges toward the values of the NOSAT experiment, making the impact practically zero. This trend is most apparent in Figure 3 in which East China was randomly chosen to depict changes in S1 skill score for the 24, 48 and 72 hours sea level pressure forecasts. The skill scores range from <60 at 24 hours to 80 at 48 hours and around 90 at 72 hours. Moreover, the impacts appear to be essentially zero with the exception of ALLSAT where the impact is negative. Also, the skill score at 48 hours (˜80) is significantly larger than the typical verification statistics (˜68) shown in Table 1.

Since satellite data were introduced basically over the oceans, the impact of satellite data could potentially be different

between land and ocean. Fig. 4 presents the skill scores of 24 hour SLP forecasts for all experiments separately for land and ocean. Here, land comprises Europe, West China, East China and North America while ocean includes regions of the North Atlantic, East Pacific and West Pacific. One can readily see that over land areas the skill scores are better (~55), but with little or no positive impact. The skill scores over ocean regions (~58) also indicate little impact. Overall, the ALLSAT experiment has the worst skill score (~60). The skill score for the 48 hour forecast displayed the same pattern that was characteristic of the 24 hour results, i.e., better skill scores for land and somewhat worse ones for ocean regions (~75) with no impact. There are practically no differences in skill scores for ocean or land at 72 hours. All experiments converge uniformly to a skill score of 87.

The comparisons of land and ocean skill scores discussed above demonstrate that in general S1 skill scores are better over land than over oceans. It is possible that the poorer verification statistics over ocean are due in part to the NOSAT analyses over ocean being less accurate than those over land due to fewer observations.

Skill scores were also computed for 500 mb height. Fig. 5 shows the 24 hour 500 mb height skill score for all experiments at different verification locations. The two adjacent locations, West pacific and East China, display the best skill score (~20) with the exception of the ALLSAT experiment. The impact however appears to be negligible. Overall the SAT experiment over the North Atlantic

shows the most improvement (-11% change) over the NOSAT experiment; the ALLSAT, the least for all regions. As was the case for SLP, the skill scores for the 48 hour 500 mb height forecasts increase according to the pattern shown in Figure 5. At 72 hours all experiments converge to the same skill scores for a given verification area with the exception of the West pacific and Eastern China where ALLSAT was still the poorest.

The skill scores of the 500 mb height forecasts were computed for land and ocean to allow comparisons. In general, the skill scores are lower (i.e. better) over land than over oceans, similar to the case of S1 for SLP. The skill score increases from around 30 at 24 hours to around 40 at 48 hours, and then to around 50 at 72 hours. Interestingly, ALLSAT always has the highest S1 scores for all three forecast periods. Again these results failed to yield any impact (no figure shown).

5.2 Root-Mean Square Differences

Another measure, the Root-Mean-Square difference was picked to verify the impact of satellite data. Bear in mind, however, that S1 skill scores and RMSE may not give the same sign to the impact of satellite data on a forecast because of their basic differences in nature. Fig. 6 shows the RMSE for the 24 hour forecast of sea level pressure which is distinctively different from Fig. 2 which depicts the 24 hour S1 skill score for SLP. While three regions, including East pacific, Europe and West China all exhibit a large RMSE in Fig. 6, the S1 skill score distribution in Fig. 2 is decisively different. As far as the impact goes, the

most noticeable improvement is over the West pacific where SAT and SATGUESS both appeared to exhibit about a 6% decrease in RMSE compared to the NOSAT experiment. This improvement, however, does not follow a predictable pattern for the 48 and 72 hour forecasts. At 48 hours the East pacific and Europe were the only two regions where SAT and ALLSAT showed some positive impact. At 72 hours, satellite data has virtually no consequence except over Europe and West China where ALLSAT seems somewhat better.

A somewhat surprising distribution of 24 hour SLP forecast RMSE is shown in Fig. 7. Contrary to the larger skill score over the ocean that we observed in Fig. 4, the errors shown in Fig. 7 are smaller over ocean than land. This is also in contrast with findings by Thomasell et al., (1986) that errors were larger over ocean than land for 48 hour forecasts. In this study, however, no difference in error was found between ocean and land at 48 hours. This result points to the fact that the satellite soundings that were assimilated into the model produce slight impact, at best, only at 24 hours and hardly any impact at all beyond that period. This holds true for both S1 skill scores and RMSE in the case of SLP.

RMS errors of 500 mb height forecasts for 24, 48 and 72 hours were also examined. Fig. 8 shows the 24 hour 500 mb height RMS error. East China stands out showing the smallest RMS error among all the regions examined. The impact however is negative, particularly in the case of ALLSAT which shows the most pronounced errors from all the experiments. At 48 hours, the RMS errors range

from the fifties to the nineties, varying widely with region. The only positive impact was seen over the North Atlantic region with about a 10% improvement for the SAT experiment.

RMS errors of the 500 mb height forecasts demonstrate a substantial difference between ocean and land. For example, at 48 hours (Fig. 9) land and ocean errors differ by more than 15%. The improvements over both land and ocean however are negligible, although negative impact appears to be rather significant.

5.3 Systematic Errors

Before detailing our findings regarding the BMC forecast model's systematic error, we will, for the reader's convenience, list the major findings below.

1) The BMC model usually underforecasts the strength of features in the sea level pressure (SLP) field, i.e., pressures are too low near high pressure systems and too high near low pressure systems.

2) The nature of the systematic errors found in the 500 mb height forecasts is not as clear cut as that of the SLP forecasts, but most often the same type of pattern is seen, i.e., the heights in troughs are not low enough and those in ridges are not high enough.

3) The use of satellite data in the BMC analysis/forecast system is found to have an impact upon the model's forecasts of sea level pressure and 500 mb height. Systematic errors in the vicinity of surface lows/500 mb troughs over the oceans were usually found to be significantly reduced. A less conclusive mix

of positive and negative impact was found for all other types of features.

5.4 Systematic Error of Sea Level Pressure Forecasts

Systematic errors of the 24, 48 and 72 h sea level pressure forecasts produced by the BMC forecasts model have been computed, according to equation (2), as the difference between the mean forecast and mean verifying (NOSAT) analysis. Fig. 10 shows the mean verifying (mean NOSAT) sea level pressure analysis for the period 9-31 January 1979 at 1200 GMT. The mean SAT, ALLSAT and SATGUESS sea level pressure analyses are essentially the same as the NOSAT and therefore are not presented. Figs. 11a-d show the mean 48 h forecast fields for this same period for the NOSAT, SAT, ALLSAT and SATGUESS experiments, while Figs. 12a-d show the corresponding systematic error fields.

In analyzing the nature of the BMC forecast model's systematic error we will focus on five features of the mean NOSAT sea level pressure analysis shown in Fig. 10. These are as follows:

1) Siberian High. The central pressure is 1044 mb and it is centered near 47N, 85E, with strong ridging (>1024 mb) extending over the North Pole.

2) Aleutian Low. The central pressure is 989 mb, with the center near 55N, 180E. An especially pronounced trough extends from the low center along the Alaskan and W. Canadian coasts.

3) Eastern Pacific High. The central pressure is 1024 mb, located near 33N, 150W.

4) Newfoundland Low. The central pressure is 1005 mb, located near 45N, 55W.

5) British Isles Low. The central pressure is 1007 mb, located near 55N, 5W. A sharp trough extends to its northeast, curving along the Norwegian coast, through the Barents Sea and into the north central Soviet Union.

The computed systematic errors of the NOSAT, SAT, ALLSAT and SATGUESS 48 h forecasts of sea level pressure are shown in Figs. 12a-d. One cannot assume that these errors accurately represent the climatology of the BMC forecast model's systematic error since they are only associated with one 25 day data period (Jan. 6-13, 1979). However, some features in the error fields are familiar from other northern hemisphere forecast models whose systematic errors have been described in the literature. These and other features arising from verification using the mean NOSAT sea level pressure analysis shown in Fig. 10 are described below by geographic region.

a. Eurasia

A large region of negative error associated with the Siberian High is seen in Figs. 12a-d, with maximum values of -16, -16, -14 and -16 mb respectively. Figs. 11a-d show that all four model experiments accurately forecast the mean position of the high's center, but are too weak in forecasting the central pressure by about -10 mb. This negative error in predicting the Siberian High's strength is common to many models, including the U.S. NMC 7-LPE model (Wallace and Woessner (1982), Bettge (1982)) and the

IMS 5-LPE model (Wolfson et al., 1985).

The maximum negative errors for all experiments are found near 50N, 50E. These are associated with the mean ridge/trough positions north of the Black and Caspian Seas. As we see in Figs. 12a-d, none of the 48 h SLP forecasts show strong ridging extending northwestward into Scandinavia as in the mean NOSAT analysis (Fig. 10). Instead, they develop an eastward protruding trough extending from Scandinavia to the north central Soviet Union. Fig. 10 shows troughing of this extent having its axis much further to the north, through the Barent's Sea.

In association with the mean forecasts' lack of ridging into Northern Europe, we find that the area of low pressure centered over the British Isles in Fig. 10, is, by all four experiments, forecast further to the east, near the Baltic Sea (Figs. 11a-d). The forecast of its central pressure is lower by only 1-3 mb, but as noted above the axis of the trough which extends eastward is much too far to the south. The nature of these features combines to add to the large area of negative error extending from the northwestern Soviet Union into northern Europe.

All four 48 h mean SLP forecasts (Figs. 11a-d) do a fairly good job of depicting the ridging extending from the Siberian High over the North Pole. However, they all develop a pronounced bogus ridge from near the center of the high well out into the western Pacific. This results in positive errors of 10 mb just east of Japan for all forecasts, as seen in Figs. 12a-d. A notable region of negative error (-12 mb for all four experiments) is also seen

over eastern Siberia in Figs. 12a-d. This is mainly due to the Aleutian Low being forecast too close to the Soviet coast, as will be discussed in the next section.

b. Pacific Ocean

A pronounced region of positive error associated with the Aleutian Low is seen in Figs. 12a-d, with maximum values of +11, +9, +10 and +10mb respectively. Figs. 11a-d show that the cause of these errors is twofold. The forecasts of the low's central pressure are too high by +4 to +7 mb and all forecasts have the low centered too close to the Soviet coast instead of further out in the Bering Sea as shown in Fig. 10. this shifting causes the largest errors to occur along and to the south of the Aleutian Islands. Figs. 1b and 3b of Wolfson, et al. (1985) show the IMS 5 LPE model (which used a subset of the observational dataset used in this study) treating this low in a remarkably similar manner.

Figs. 12a-d show regions of negative error in the east Pacific; the most pronounced error (-6 mb) being in the NOSAT forecast (Fig. 12a). The regions of largest error are found to the north of the east Pacific High's center, illustrating that much of the error is simply due to the forecast high not being broad enough, as the location of the main enter is forecast quite well in all experiments. The central pressures in the forecasts are low by only -2 to -4 mb.

c. West Atlantic

Figs. 11a-d indicate that all four experiments have major difficulties in forecasting the occurrence of the mean low in the

northwest Atlantic. The mean NOSAT analysis (Fig. 10) shows a 1005 mb low located off the coast of Newfoundland and Nova Scotia. This feature is barely hinted in the mean 48 h forecasts (Figs. 11a-d), resulting in large areas of positive error in Figs. 12a-d with maximum values of +12, +11, +9 and +12 mb respectively. Contributing further to these areas of positive error are the experiments' tendencies to develop a subtropical high off the southeast coast of the U.S. (Figs. 11a-d). The mean NOSAT analysis (Fig. 10) does show a broad area of relatively high pressure to the south of the Newfoundland Low, but it is weak (maximum pressure 1018 mb) and is centered much farther to the east than the forecast highs.

The systematic errors described above indicate (over this data period at least) that the BMC forecast model usually produces positive errors (SLP too high) in the vicinity of surface low pressure systems. Negative errors (SLP too low) are usually found in the vicinity of surface highs.

5.5 Systematic Error of 500 mb Height Forecasts

Systematic errors of the 24, 48 and 72 h 500 mb height forecasts produced by the BMC forecast model have been computed in the same manner as described in section 5.4 for the sea level pressure. Figs. 13a-d show the mean NOSAT, SAT, ALLSAT and SATGUESS 500 mb analyses for the period 9-31 January 1979 at 1200 GMT. Significant differences between these analyses are difficult to spot when visually comparing them, therefore difference maps which compare analyses using satellite data to the NOSAT analysis

(Z_s (satellite experiment) - Z_{ns} (NOSAT) are presented in Figs. 14a-c.

The inclusion of satellite data over the oceans in the SAT and SATGUESS experiments is responsible for the differences in the height field shown in Figs. 14a and 14c. Additional height differences over land in the ALLSAT difference diagram (Fig. 14b) are due to exclusive use of satellite data in place of radiosonde observations over the entire model domain. Such differences are often associated with regions of high surface elevation, e.g., the Greenland icecap, where satellite soundings typically are in error.

The mean 48 h 500 mb height forecasts for the four experiments are shown in Figs. 15a-d and the corresponding systematic error fields are shown in Figs. 16a-d. Descriptions of the main features in the systematic error fields are presented below by geographic region.

a. Eurasia

Fig. 16a-d all show a broad region of negative error centered near the axis of the ridge located near 50N, 40E in the NOSAT analysis (Fig. 13a). The largest height errors for the four experiments are -88, -81, -86 and -87 meters, respectively. This region of negative error extends westward to cover all of Europe, coincident with the troughing that is seen in all mean analyses (Figs. 13a-d) and 48 h forecasts (Figs. 14a-d). The nature of the forecast errors for this trough/ridge pattern is consistent with that of the associated regions of low and high sea level pressure described in section 3.1, i.e., the strength of the trough, like

the surface low, is overforecast and that of the ridge, like the surface high is underforecast.

Figs. 15a-d all show a region of positive forecast error over eastern China, with the largest errors being +39, +46, +61 and +40 meters, respectively. These seem to be caused by the mean 48 h forecasts not predicting the slight troughing that is evident over eastern China in the mean NOSAT analysis (Fig. 13a). The mean ALLSAT 48 h forecast (Fig. 15c) does show troughing a bit farther inland but actually builds a slight ridge over the Yellow Sea near 35N, 120E, contributing to the +61 m error there.

b. Pacific Ocean

In the Pacific, north of about 20 degrees latitude, there is a distinct pattern of mean forecast error common to all four experiments. in figs. 16a-d we find a region of negative error in the west Pacific (near 30N, 150E), an area of positive error in the north central Pacific (near 45N, 175E) and another region of negative error in the east Pacific (near 45N, 140W).

The mean NOSAT analysis (Fig. 13a) shows fairly zonal flow in the western Pacific (between 140E and 160E), whereas all mean 48 h forecasts show a bit of troughing, thus contributing to the negative error found there. The trough extending into the northern Pacific in the NOSAT analysis is quite rounded with its main axis lying roughly to the south of the Aleutian Islands. All mean 48 h forecasts sharpen this trough, rotating its main axis counterclockwise into the Bering Sea. This results in increased heights to the south of the Aleutians, leading to the region of

positive error seen there.

All experiments do a fair job of predicting the location and mean height at the center of the closed low over the northeastern Soviet Union. Maximum errors here are only on the order of -10 m.

The position of the east Pacific ridge shown in the NOSAT analysis is forecast well by all four experiments but all underforecast its strength, resulting in maximum errors of -62, -47, and -66 meters, respectively. Interestingly, the strength of the extension of this ridge into polar regions is overforecast in all experiments by +50 to +60 meters.

c. North America

Figs. 16a-d all show a large region of positive error extending from polar regions, down through central Canada and into the southern and eastern U.S. Mean 48 h forecasts from each experiment (Figs. 15a-d) all show two regions of maximum error; one centered over the Northwestern Territories of Canada and the other over the northeastern United States. Maximum errors range from +60 to nearly +80 meters. All experiments do a good job of positioning the axes of the troughs found in this large region, but their forecasts of the heights are all too high.

d. Atlantic Ocean

There are two features of interest in this area. One is the ridge in the north Atlantic, whose axis is seen to lie roughly along 35 to 40W in the mean NOSAT analysis (Fig. 13a). As has been the case with other features discussed, the mean position of the ridge is forecast quite well by all four experiments, but its

strength is not. All experiments underforecast the heights here, with the greatest errors being -60, -58, -74 and -64 meters, respectively.

The other feature leading to some significant errors is the trough just off the west coast of north Africa. It appears that the SAT and ALLSAT experiments, while positioning the axis of the trough accurately, have deepened it too much, resulting in maximum errors of -60 and -53 meters, respectively.

Interpretations of the BMC model's 500 mb height forecasts are not as straight forward as those made from the sea level pressure forecasts. In all experiments described here, the model often shows positive error in the area of troughs and negative error in the vicinity of ridges; however this is not always the case. The mean heights in the troughs over Europe and off the northwest African coast were often forecast too low while those near the crest of the east Pacific ridge were forecast too high. The mean height of the one closed low found in the averaged field was predicted quite accurately in all experiments.

5.6 Differences in Systematic Errors of Sea Level Pressure Forecasts

To determine any differences between systematic errors from the experimental runs which used satellite data and the NOSAT experiment, the difference between the absolute values of the systematic errors from the 48 h mean SLP forecasts were analyzed. Figs. 17a-c show isopleths of E_s-E_{ns}, where E is the systematic error, S denotes a satellite data experiment (SAT, ALLSAT or

SATGUESS) and NS denotes the NOSAT experiment. Negative values imply a reduction in systematic error, positive values an increase in systematic error. Heavy dashed lines show areas where E_s is significantly different from E_{ns}.

a. SAT minus NOSAT

Fig. 8a shows three general regions where the use of satellite data in the SAT experiment has a significant impact on the mean 48 hour forecast. All three areas of significance relate to a reduction in systematic error. These regions all have in common their relative location to the southeast of surface low pressure systems seen in the mean NOSAT analysis (Fig. 10). The region of reduced error found to the southeast of the mean low analyzed over the British Isles is especially expansive, stretching from central Europe to the Middle East.

b. ALLSAT minus NOSAT

The ALLSAT experimental run shows many areas where satellite data has significant impact (Fig. 17b). Most of the impact leads to reductions in systematic error. Two such areas, those in the western Atlantic and northern Pacific, are similar to those found in the case of SAT minus NOSAT (Fig. 17a).

A very large area of significantly reduced error stretches from western Europe to central Asia and another smaller area is seen in the northeast Soviet Union. Areas of significantly increased error are seen in eastern China and north and south of Japan (Fig. 17b).

c. SATGUESS minus NOSAT

The use of satellite data in the SATGUESS experiment produced no significant impact.

These experimental results indicate that satellite data have an impact on the pattern of the systematic error of the BMC model's sea level pressure forecasts. The use of satellite data over the oceans seems to have, for the most part, a beneficial effect upon the sea level pressure forecast in the vicinity of well developed oceanic surface lows. For this data period at least, it also seems to reduce forecast error over western Eurasia, where a relatively flat mean sea level pressure field is observed.

The additional use of satellite data over land caused eastward expansion of the area of beneficial impact that was seen over western Eurasia in the Fig. 17a. However, in east Asia areas of significant impact mainly showed increases in systematic error. Similar patterns of impact were generally not observed over North America for either type of satellite data usage.

5.7 Differences in the Systematic Error of 500 mb Height Forecasts

Differences in the systematic error of 500 mb height forecasts were calculated according to the method described in section 5.6 for sea level pressure forecasts. The resulting fields are presented in Figs. 18a-c. As before, negative values imply a reduction in systematic error, positive values an increase in systematic error, and heavy dashed lines denote areas where E_s is significantly different from E_{ns}.

a. SAT minus NOSAT

Fig. 18a shows several regions where the use of satellite data over the oceans has significant impact on the mean 48 h 500 mb height forecast. In the north Pacific, two fairly large areas of reduced error straddle a smaller area of increased error. The two areas of reduced error were found to be statistically significant while the area of increased area was not. These areas are mainly associated with the treatment of the deep trough shown there in the NOSAT analysis (Fig. 13a).

Another area of significant impact lies along the coasts of China and Japan. Over the ocean south of Japan there is a reduction in error, whereas just inland over China we see an increase in error. These features are related to subtle differences in treatment of the mainly zonal flow in that area that may be seen in Figs. 15a-b.

An area of increased systematic error is seen over south central Asia. This feature lies in the same area as a trough shown in the mean NOSAT analysis (Fig. 13a).

Two areas of significant impact are seen over Europe and northwest Africa in Fig. 18a. Over Europe there is a reduction in systematic error and along the west coast of north Africa there is an increase in error. The mean NOSAT analysis (Fig. 13a) shows a trough in each of these areas. The use of satellite data over the oceans evidently aided in forecasting the nature of the well developed European trough but for an undetermined reason was detrimental in forecasting the weaker trough off the African coast.

b. ALLSAT minus NOSAT

Fig. 18b shows that the use of satellite data in the ALLSAT experiment leads to some impacts that are very similar tc those seen in the Fig. 18a. Specifically, Fig. 18b shows the same patterns of impact over the north Pacific and along the coasts of China and Japan. A large area of reduced error was found over Europe as in the SAT minus NOSAT case, but in this experiment it was not found to be statistically significant. An area of significantly increased error is again seen along the northwest coast of Africa.

Fig. 18b also shows impact in the north Atlantic in the vicinity of two features seen there in the mean NOSAT analysis (Fig. 13a). These are the trough which extends from eastern Canada over the north Atlantic and the ridge in the north Atlantic. The ALLSAT experiment reduced the error in the vicinity of the trough and increased the error in the area of the ridge. It appears that the addition of satellite data over land causes some significant increases in error in area stretching from the central to northeastern Soviet Union.

c. SATGUESS minus NOSAT

The pattern of the impact which satellite data produces in the SATGUESS experiment is quite similar to that of the SAT and ALLSAT experiments, but the magnitude of the impact is slight and generally not significant.

The results presented above indicate that satellite data have an impact on the pattern of the systematic error of the BMC model's 500 mb height forecasts. The use of satellite data over the oceans

is found to generally reduce systematic errors in the vicinity of well defined oceanic troughs. It was also found to have significant effects upon 500 mb height forecasts over land, but the exact nature of these effects, i.e., whether it typically leads to reductions or increases in systematic error, is not clear from these results.

When combining satellite data over land with that over oceans, the same general reductions in systematic error are found near well defined oceanic troughs; however we see some evidence that the use of the data to this extent may increase error in the vicinity of oceanic ridges. Areas of significantly increased error seen over east Asia in Fig. 18b may suggest that the use of satellite data over land is detrimental to the BMC model's forecasts of 500 mb height over land.

6. Discussion and Conclusion

The results of January 1979 data failed to produce a consistent accurate forecast over land and ocean. While the percent change from NOSAT in skill scores were consistently lower over land both at SLP and 500 mb height the reverse was the case for RMS error. This is in direct contrast with the results obtained by Thomasell et al., (1986). The exact cause for this ambiguity is unknown although it may related to the distribution of satellite data and conventional data with respect to a specific experiment.

Overall, the experiment results show the ALLSAT to be the poorest, i.e., higher RMS forecast errors and higher S1 skill

scores. The impact of satellite data on forecasts in general however, is negligible.

The use of satellite data in the BMC analysis and forecast system does however have an impact upon the systematic error or its sea level pressure and 500 mb height forecasts. Possible reasons why this is not consistently reflected in RMS forecast errors and S1 skill scores include: 1) areas of impact did not always lie within the experimental areas shown in Fig. 1, and 2) areas of positive and negative impact were sometimes found very close to each other, thereby tending to cancel each others' effects when found in the same experimental area.

The SAT and ALLSAT experiments showed similar, significant reductions in systematic forecast error in the vicinity of well developed oceanic surface lows and their associated 500 mb troughs. While significant impact was also observed in association with oceanic highs/ridges, also with lows/troughs and highs/ridges over land, the sign of the impact upon each type of weather system, i.e., whether there was a reduction or increase in systematic error, was not found to be consistent.

The inability of satellite data to improve the forecast raises questions about the quality of the January 1979 FGGE data set. A study conducted by Susskind et al., (1984) concludes that the RMS layer mean temperature errors of the FGGE data set was 2.4 K which is higher than 2.2 K obtained by a different retrieval algorithm adapted by the Goddard Laboratory for Atmospheric Sciences (GLAS). The GLAS retrievals were reported to be significantly more accurate

in the troposphere though slightly less accurate in the stratosphere than those in the FGGE data base. In their assessment of TIROS-N satellite soundings Gruber and Watkins (1982) report that the RMS difference varies from a maximum of 3.9 K in the 1000-850 mb layer to a minimum of 1.7 K in the 500-700 mb layer, but is generally between 2.5 and 3 K throughout most of the troposphere in the mid and high latitude zones. Although the TIROS-N data examined by Gruber and Watkins (1982) differ from that of FGGE data, the same retrieval algorithm was applied to produce the temperature profiles. However, NESDIS introduced changes to their operational retrieval algorithm in 1982 (McMillin and Dean, 1982), but this new processing system has not been applied to the data in the FGGE data base.

Table 1

Typical Verification Statistics
For 48 Hr Forecasts Over Land

	NOSAT	SAT	%Change	
	RMS Error 500mb Z (m)			
GLAS(1978)	77.9	72.8	-6.5	
NMC/(1978)	65.0	63.7	-2.0	
IMS/NESDIS/(1986)	71.8	64.8	-9.7	
	S1 Skill Score 500mb Z (m)			
GLAS(1982)	39.6	37.7	-4.8	4x5 grid
	36.6	33.5	-6.9	2.5x3 grid
NMC(1978)	34.8	34.3	-1.4	2.5 grid
IMS/NESDIS(1986)	42.7	41.4	-3.0	3.8 grid at 60N
	S1 Skill Score Sea Level Pressure			
GLAS(1982)	72.7	69.5	-4.4	4x5 grid
	67.1	60.7	-9.6	2.5x3 grid
IMS/NESDIS(1986)	68.4	67.9	-0.7	3.8 grid at 60N

References

Bengtssen, L., Kanemitsu, P. Kallberg and S. Uppala, 1982: FGGE research activities at ECMWE. Bull Amer. Meteor. Soc., 63, 277-303.

Bettge, T., 1982: An analysis of recent improvement in NMC operational forecasts, ninth Conf. Weather Forecasting and Analysis, Seattle, Amer. Meteor. Soc., 79-83.

Cressman, G.P., 1959: An operational objective analysis system. Mon. Wea. Rev., 87, 367-374.

Gruber, A., and C.D. Watkins, 1982: Statistical Assessment of the Quality of TIROS-N and NOAA-6 Satellite Soundings. Mon. Wea. Rev. 110, 867-876.

Halem, M., E. Kalnay, W.B. Baker and R. Atlas, 1982: An assessment of the FGGE satellite observing system during SOP-1. Bull Amer. Meteor. Soc., 63, 407-426.

Mao, J. 1986: The study of the impact of TIROS-N satellite data on People's Republic of China's operational numerical weather analysis and forecast. Master degree thesis, National Meteorological Center, PRC.

McMillin, L.M., and C. Dean, 1982: Evaluation of a new operational technique for producing clear radiances. J. Appl. Meteorol. 21, 1005-1014.

Susskind, J., D. Reuter and M.T. Chahine, 1984: Remote Sensing of Weather and Climate Parameters from HIRS 2/MSU on TIROS-N. J. Geophys. Res. 89, 4677-4697.

Teweles, S. Jr., and H.B. Wobus, 1954: Verification of prognostic charts. Bull Amer. Meteor. Soc. 35, 455-465.

Thomasell, A. Jr., A. Gruber, H. Brodrick, N. Wolfson and Z. Alperson, 1986: The impact of satellite soundings on the Numerical Forecasts of the Israel Meteorological SErvices. Mon. Wea. Rev. 114, 1251-1262.

Tracton, M.S., A.J. Desmarais, R.J. van Harren and R.D. McPherson, 1980: The impact of satellite soundings on the National Meteorological Center's Analysis and Forecast System - The data systems test results. Mon. Wea. Rev. 109, 543-585.

Wallace, J.M. and J.K. Woessner, 1982: Analysis of forecast errors in the NMC hemispheric primitive equation model. Mon. Wea. Rev., 109, 2444-2449.

Wolfson, N., A. Thomasell, A. Gruber and G. Ohring, 1985: The Impact of Satellite Soundings Data on the Systematic Error of a Numerical Weather Prediction Model. Mon. Wea. Rev. 113, 1031-1049.

Chinese Meteorological Grid
Experimental Areas

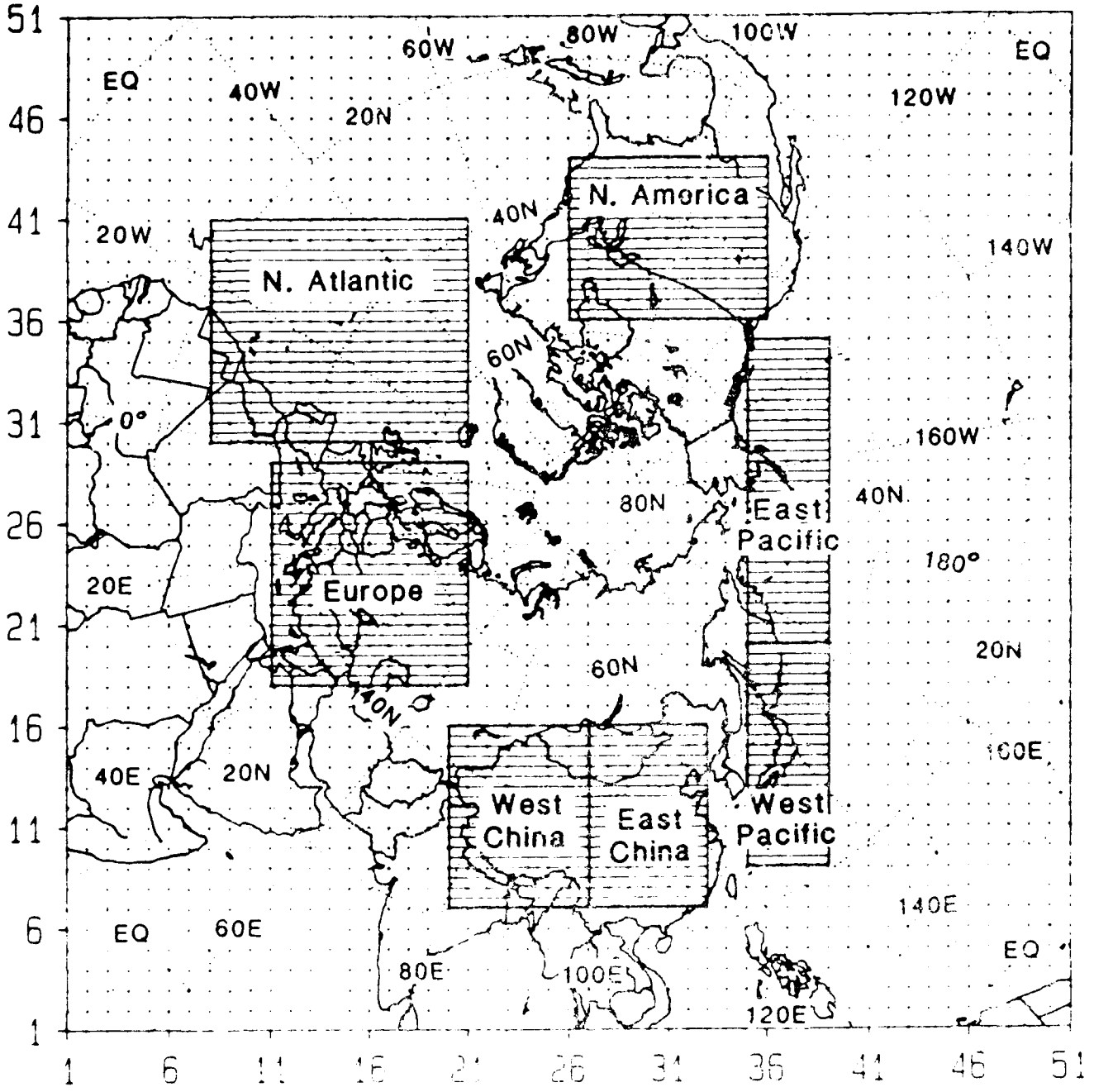

Fig. 1. Grid network of the Beijing Meteorological Center (BMC) model and seven objective verification areas.

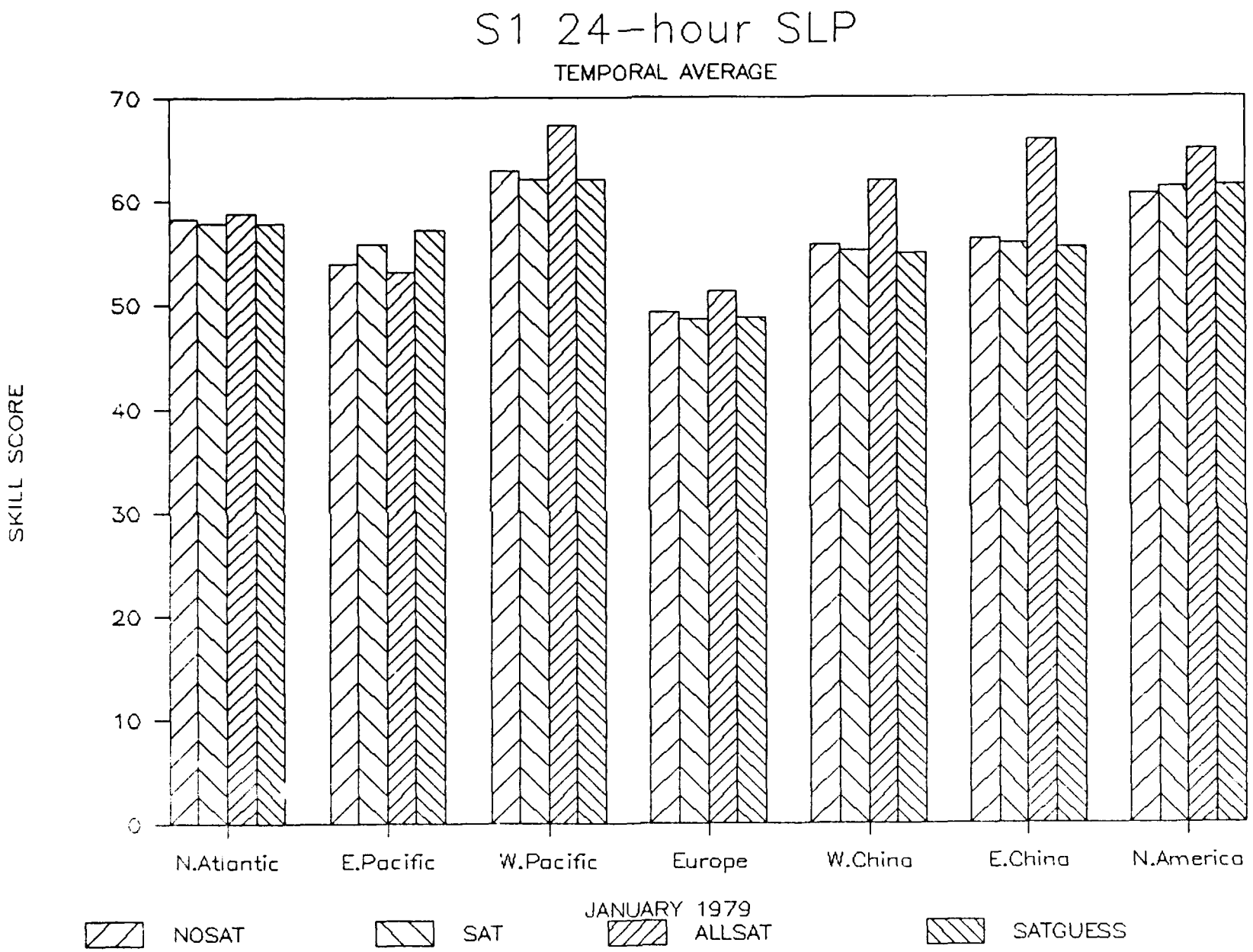

Fig. 2. Average skill score of 24-hour forecast of Sea Level Pressure (SLP) over each verification area for NOSAT, SAT, ALLSAT and SATGUESS experiments.

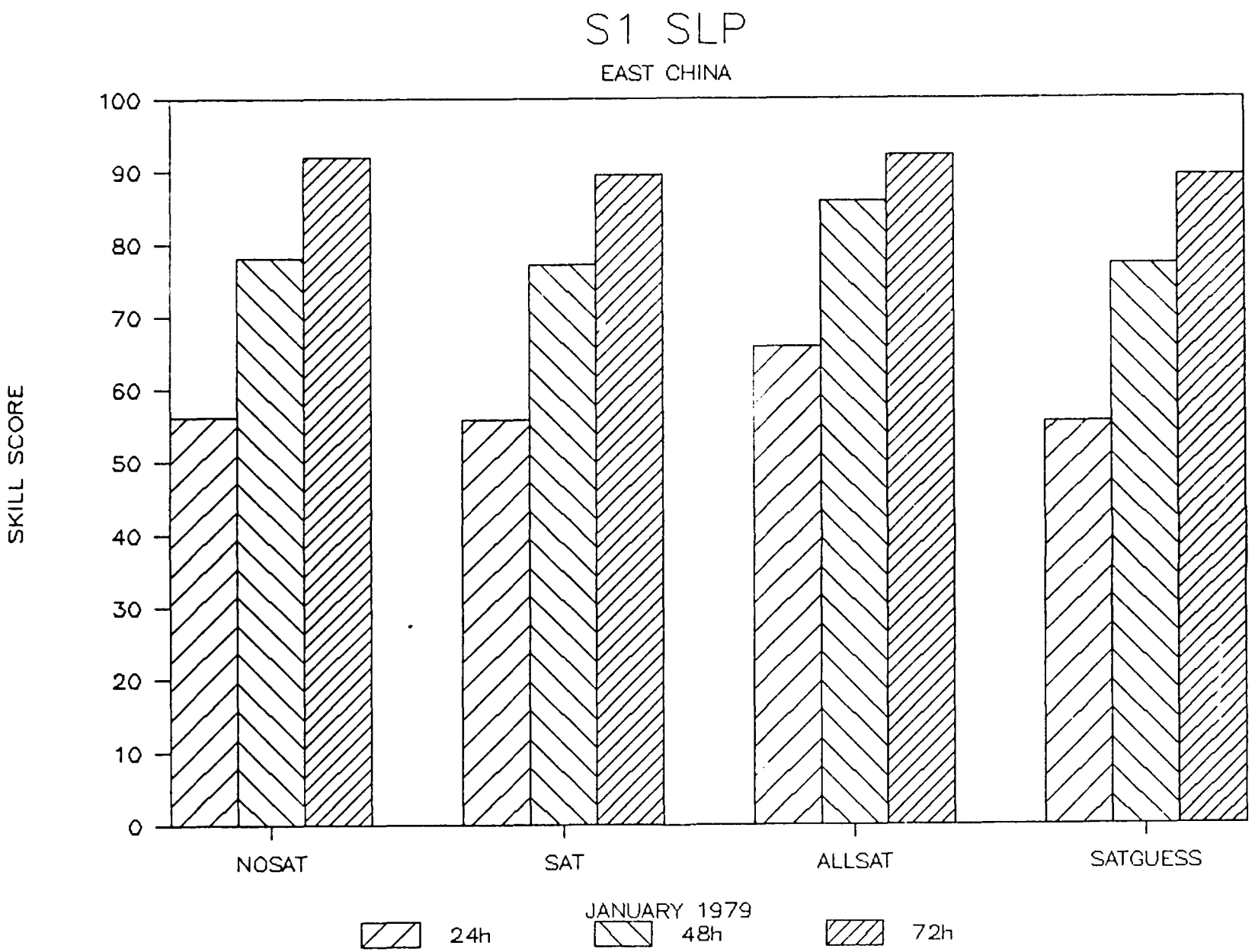

Fig. 3. Average skill score of 24, 48 and 72 hour forecast of SLP over East China for NOSAT, SAT, ALLSAT and SATGUESS experiments.

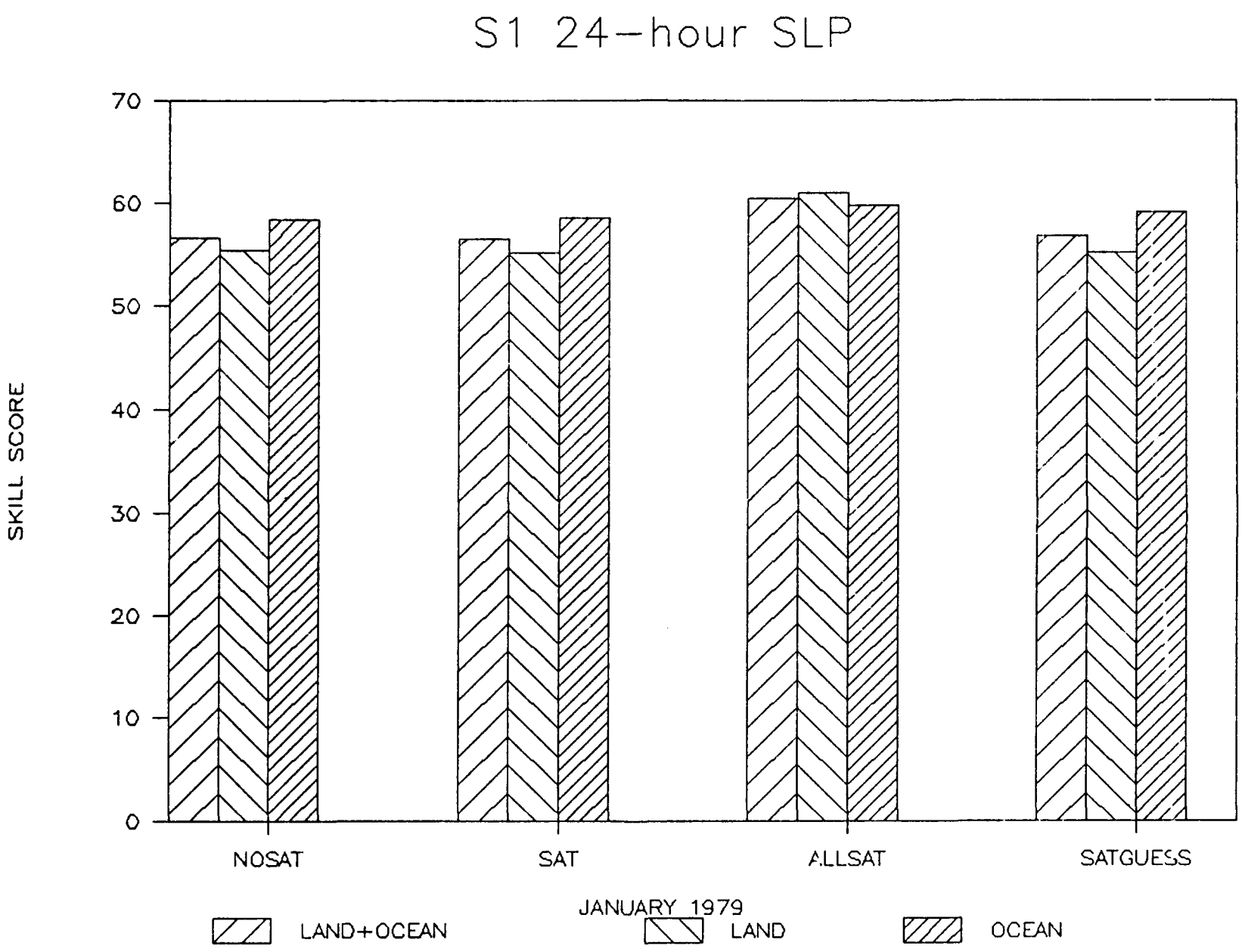

Fig. 4. Skill score of 24-hour forecast of SLP respectively over land, ocean and land plus ocean for all four experiments.

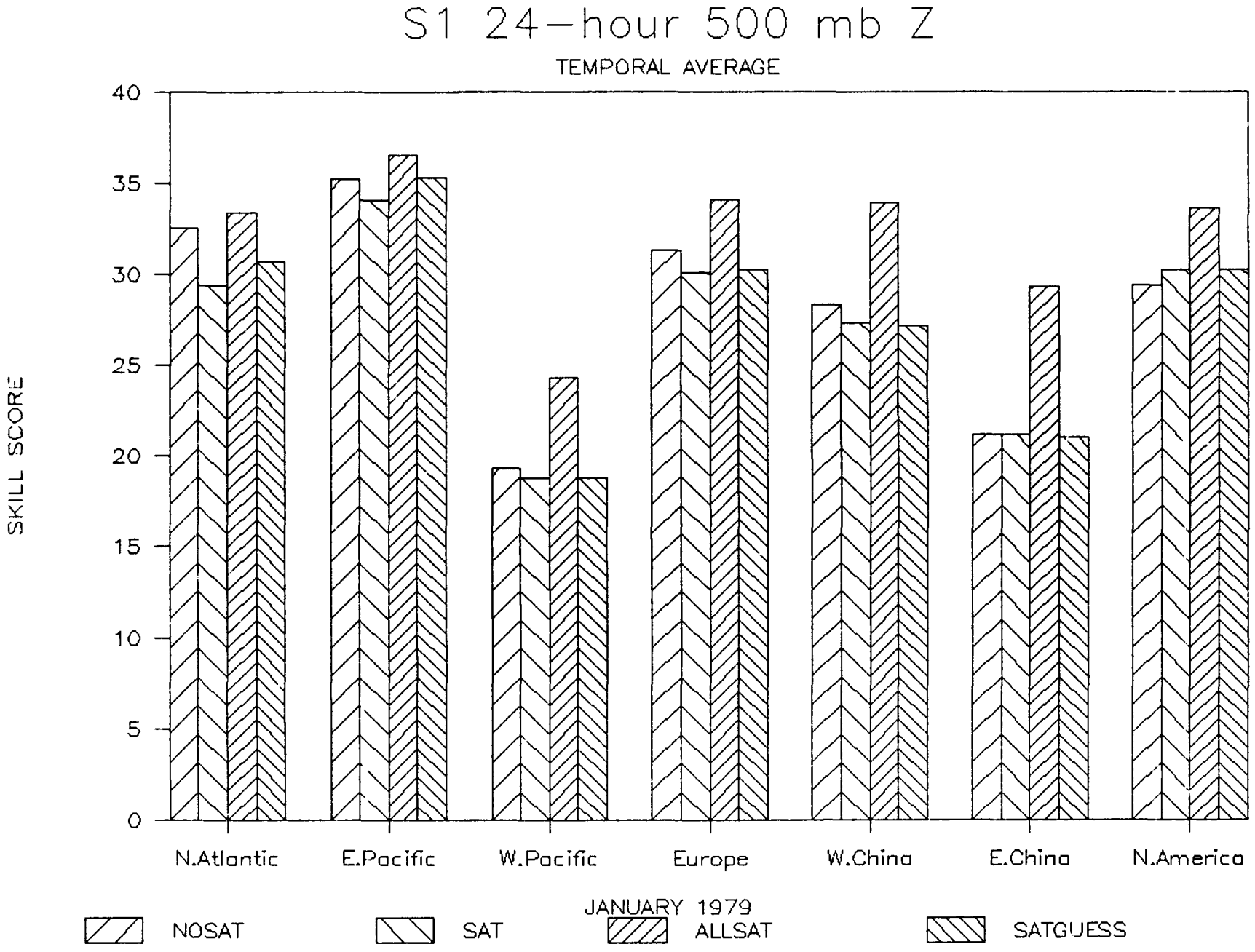

Fig. 5. Same as in Figure 2 but for 500 mb height.

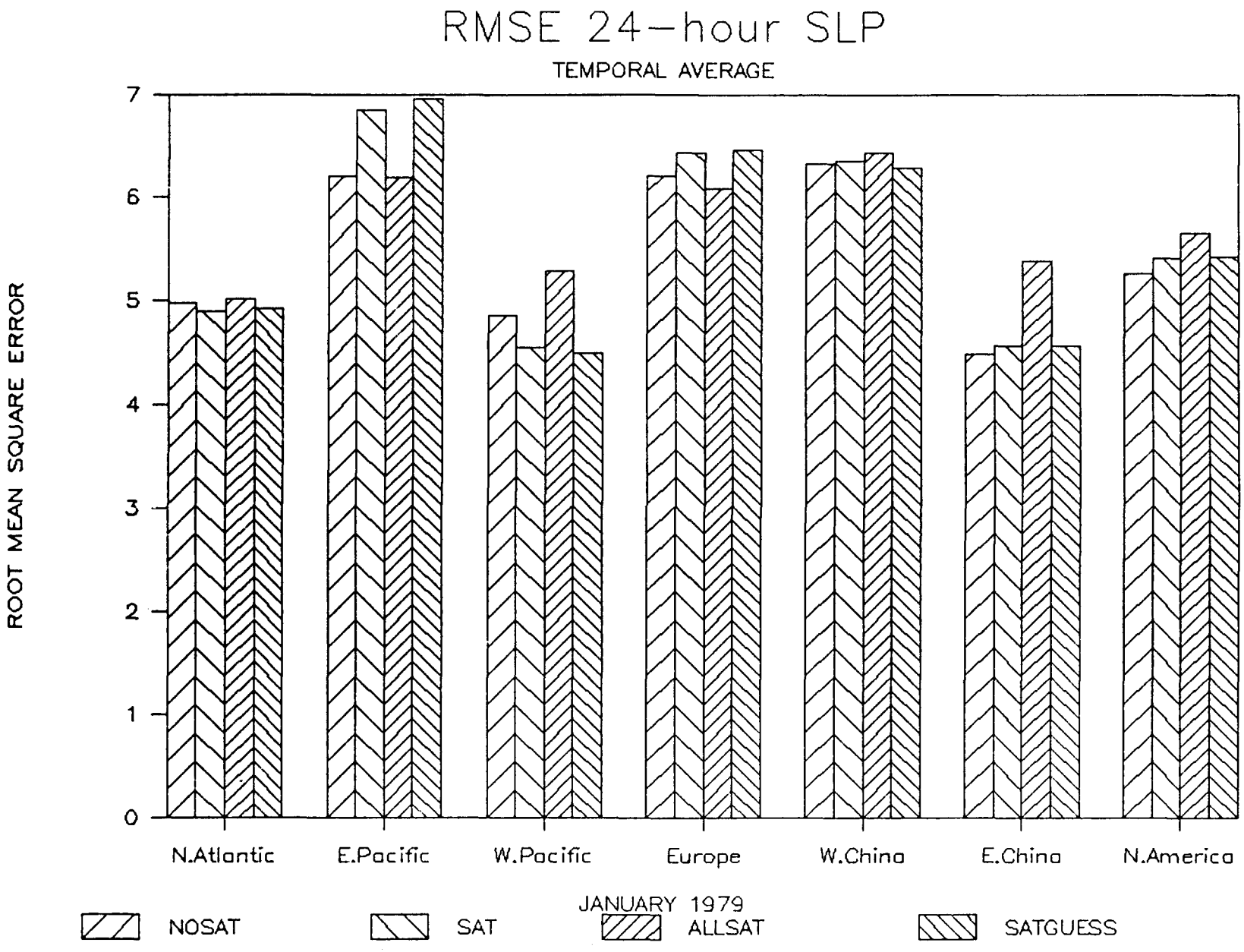

Fig. 6. Average root mean square error of 24-hour forecast of SLP over each verification area for NOSAT, SAT, ALLSAT and SATGUESS experiments.

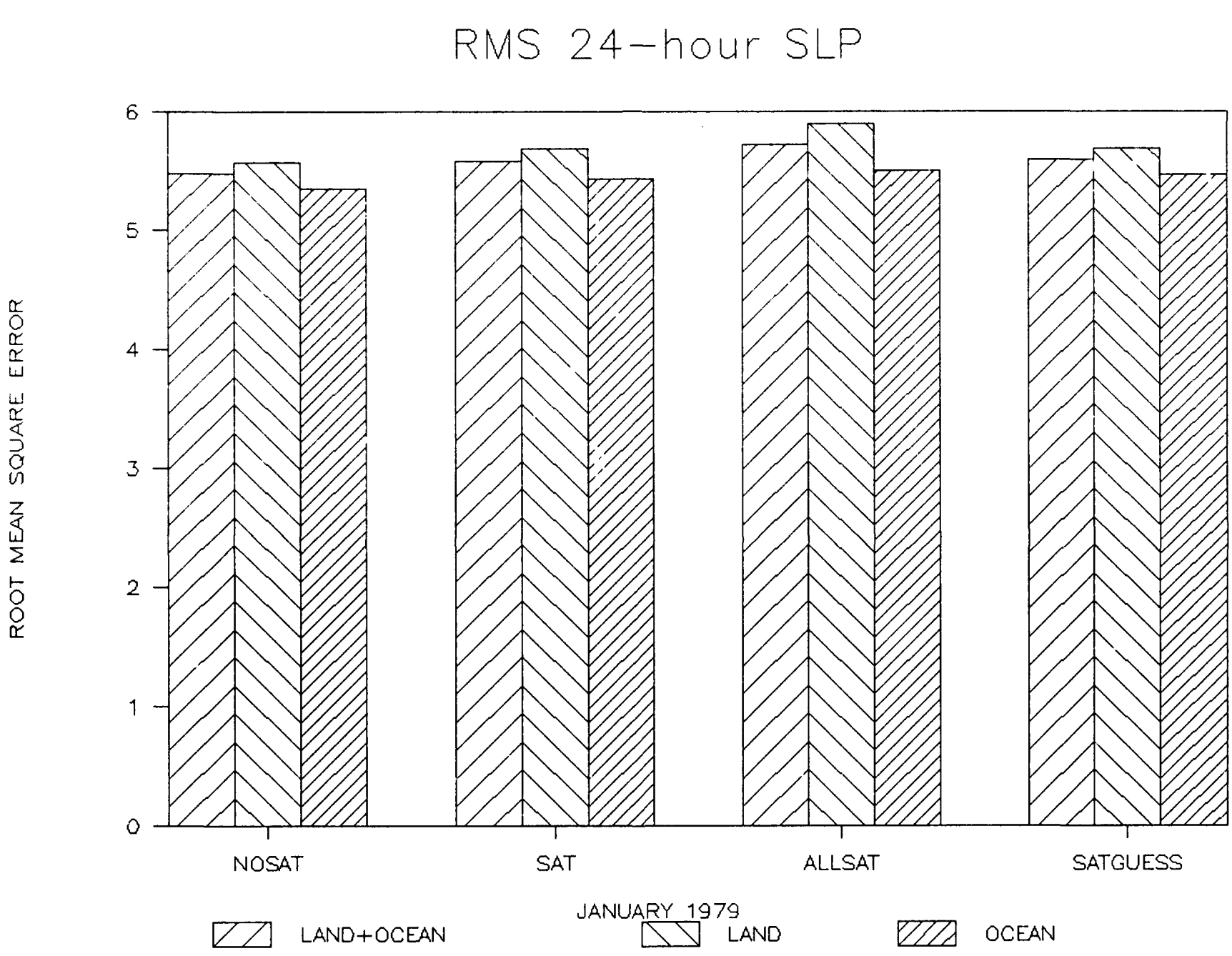

Fig. 7. Same as in Figure 4 but for root mean square error.

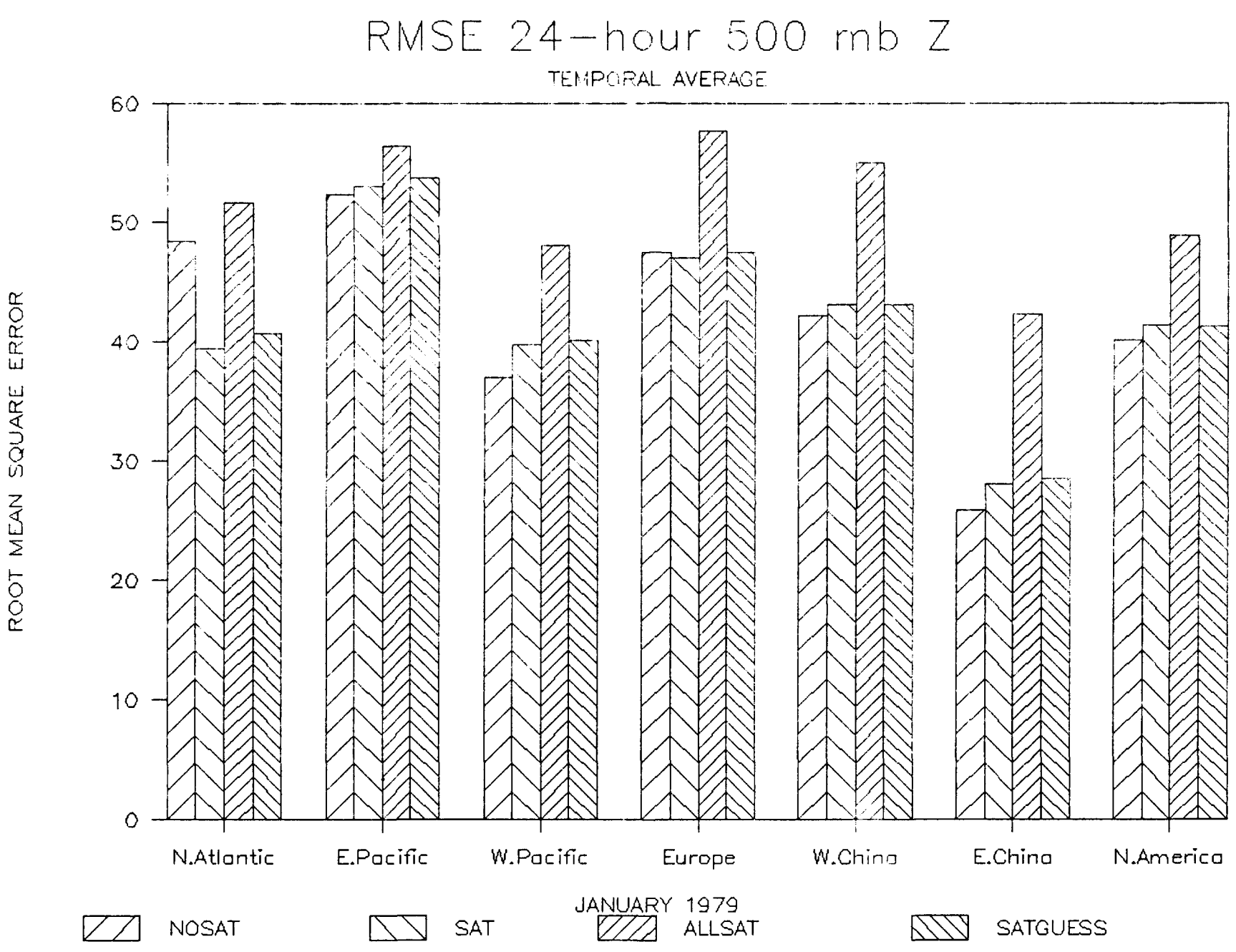

Fig. 8. Same as in Figure 6 but for 500 mb height.

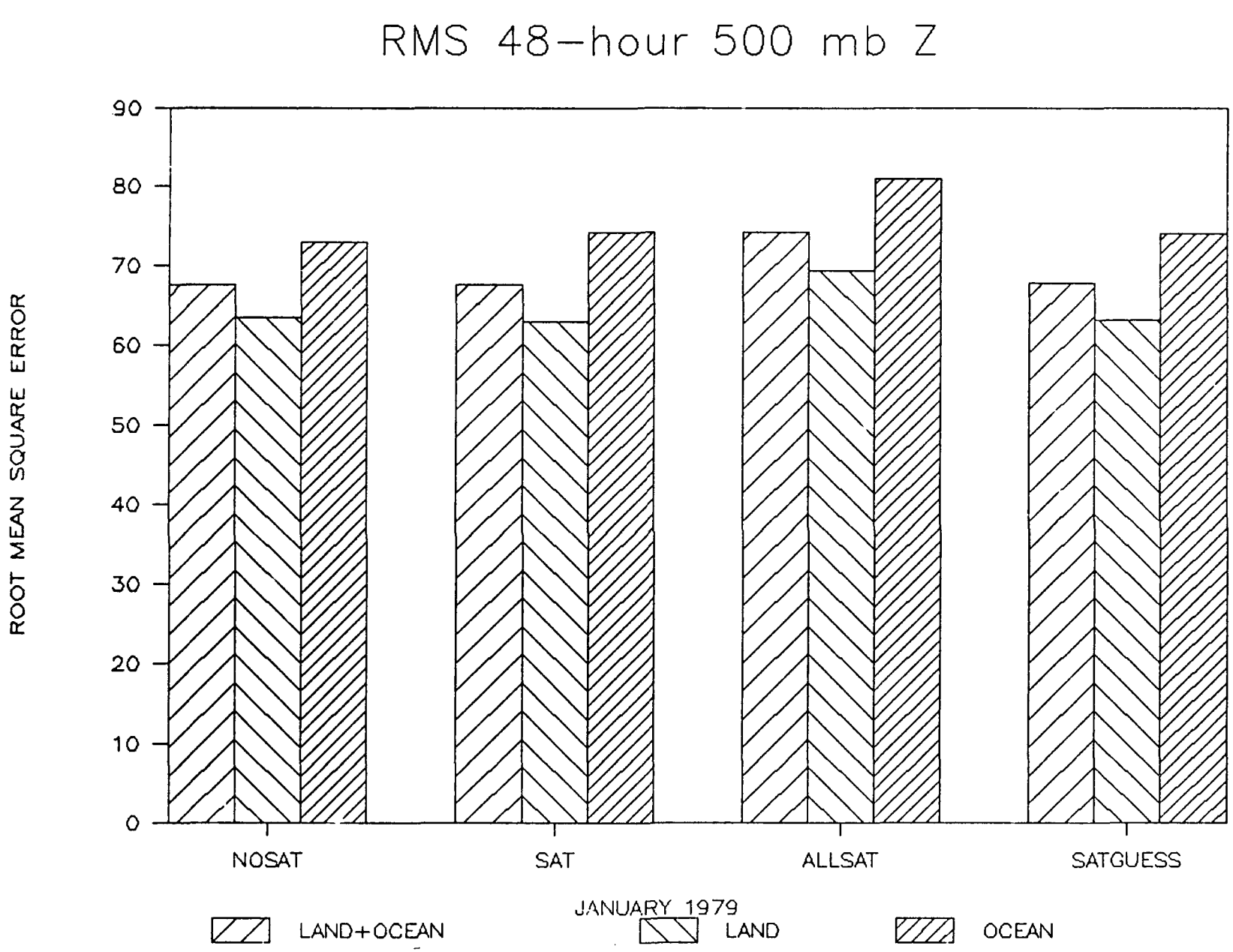

Fig. 9. Root mean square error of 48-hour forecast of 500 mb height respectively over land, ocean and land plus ocean for all four experiments.

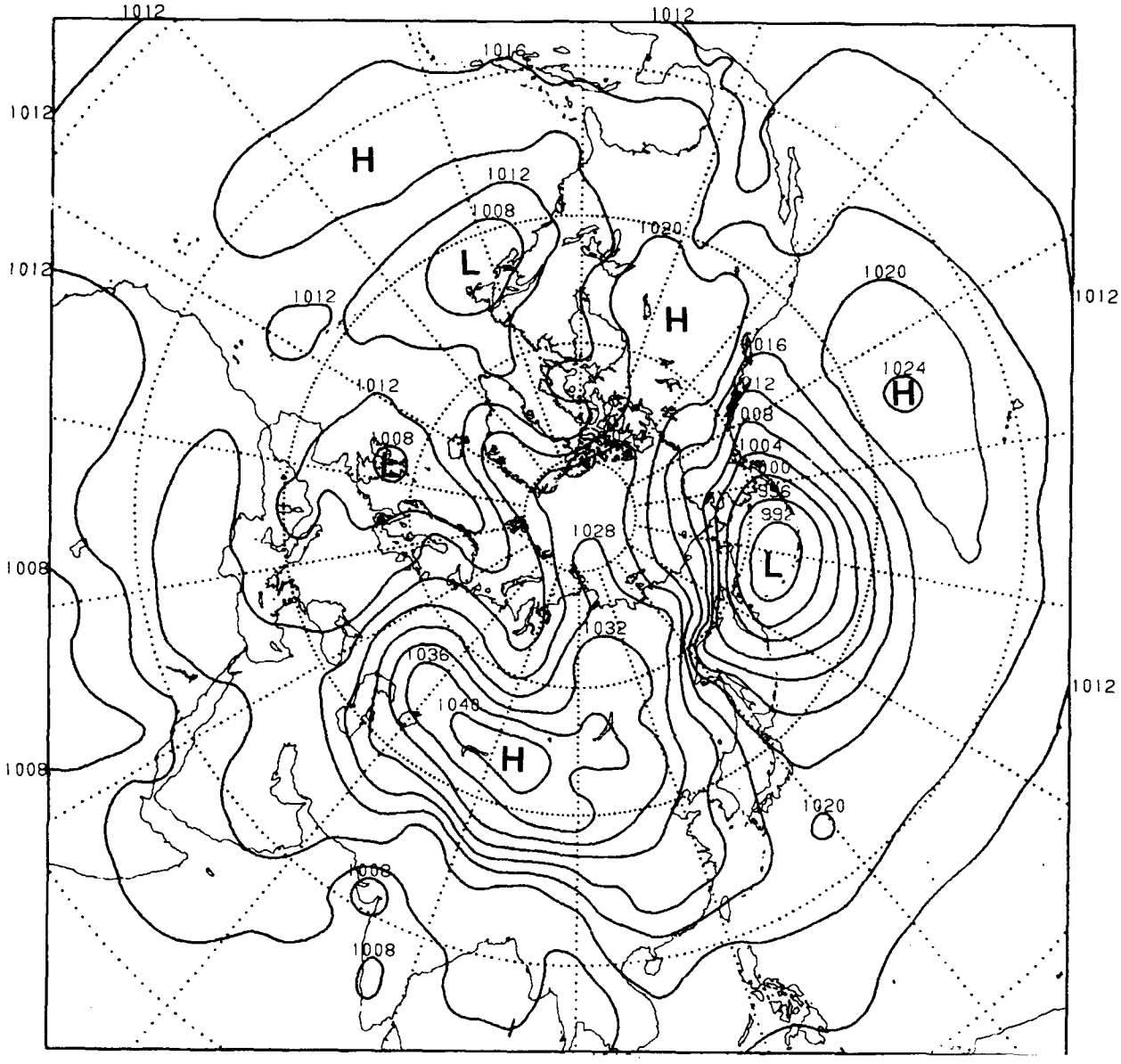

Fig. 10. Average NOSAT sea level pressure analysis for 9-31 January 1979 at 1200 GMT.

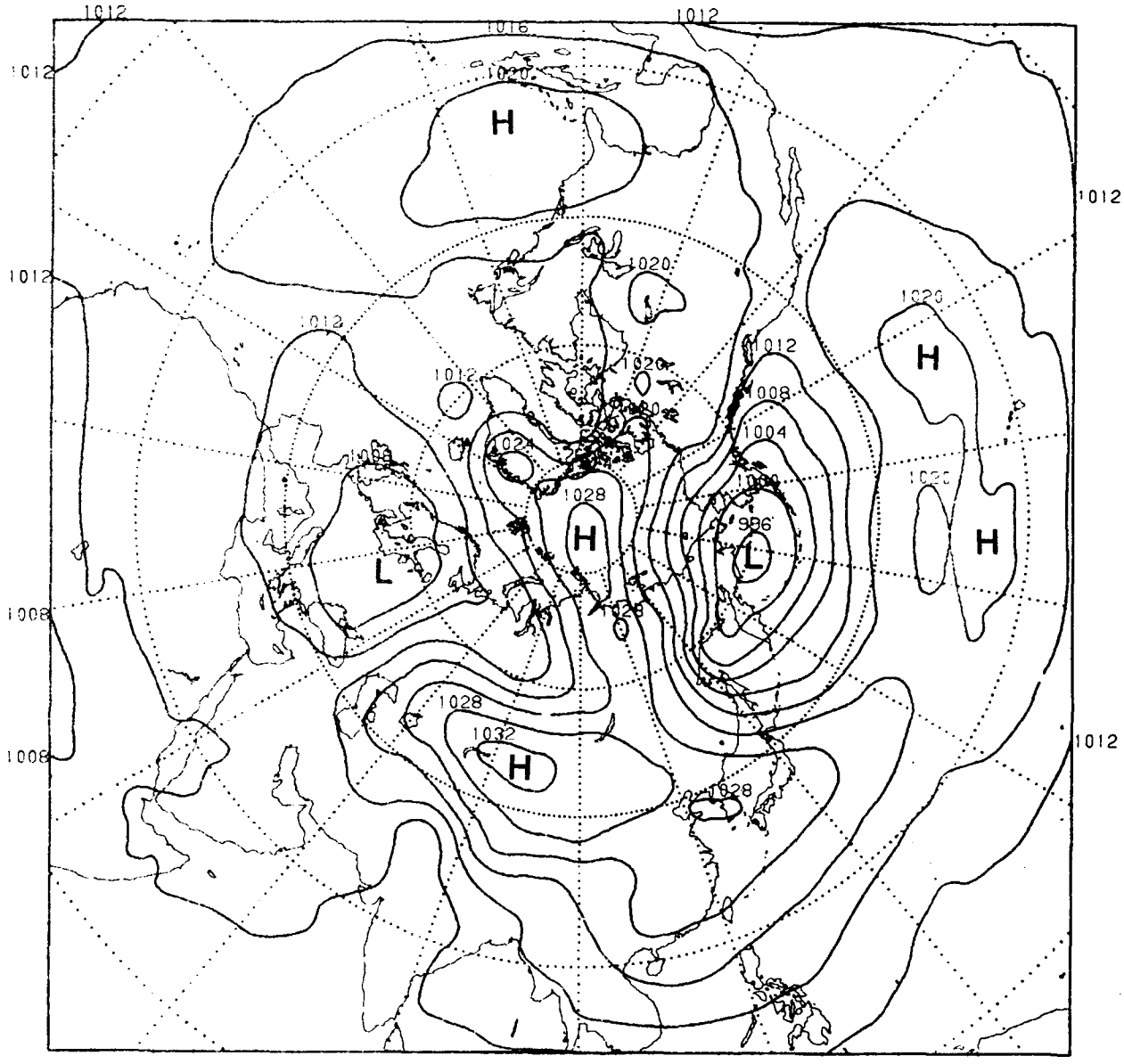

Fig. 11a. Average 48 hour NOSAT forecast of sea level pressure for 9-31 January 1979 at 1200 GMT.

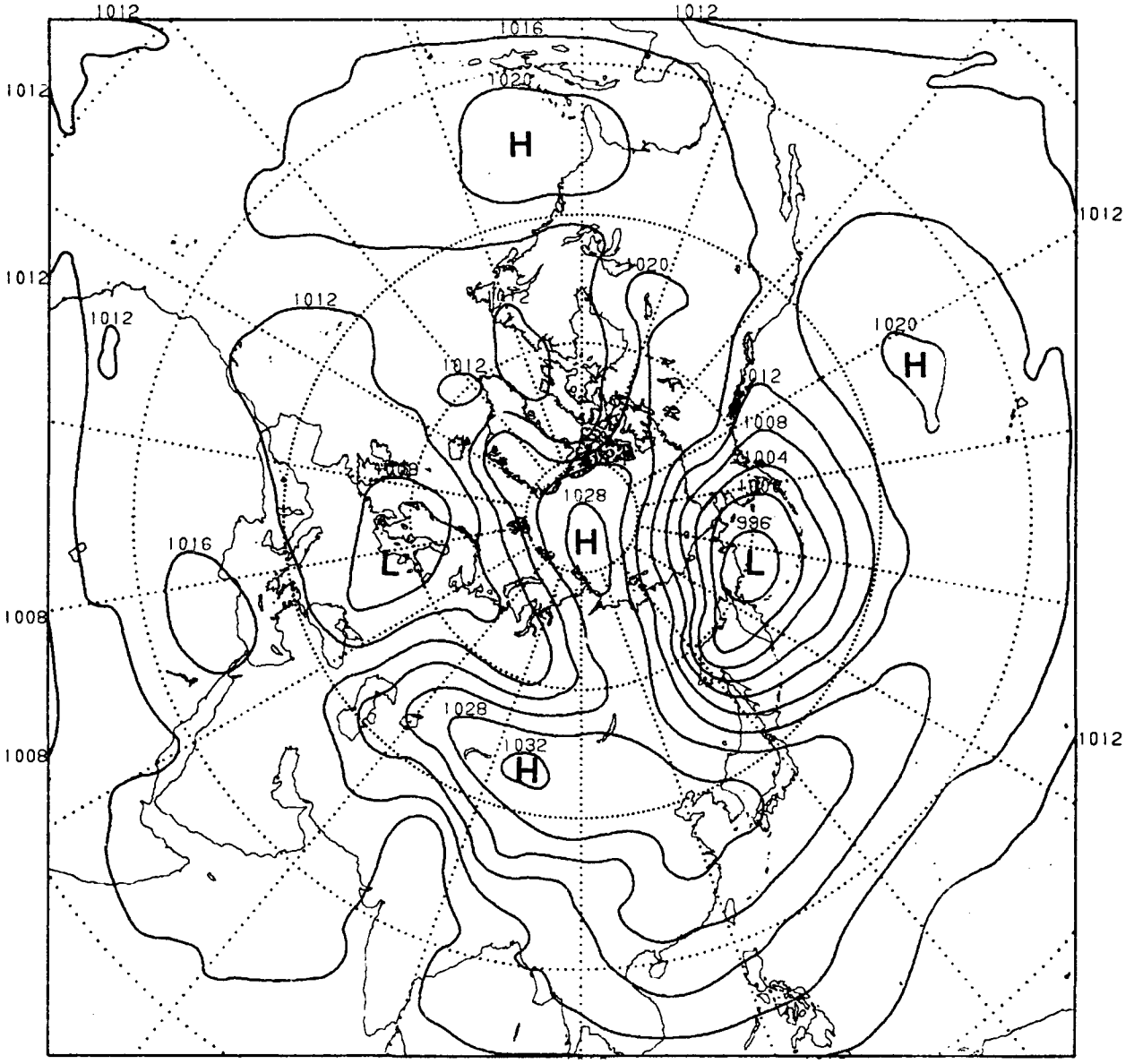

Fig. 11b. Average 48 hour SAT forecast of sea level pressure for 9-31 January 1979 at 1200 GMT.

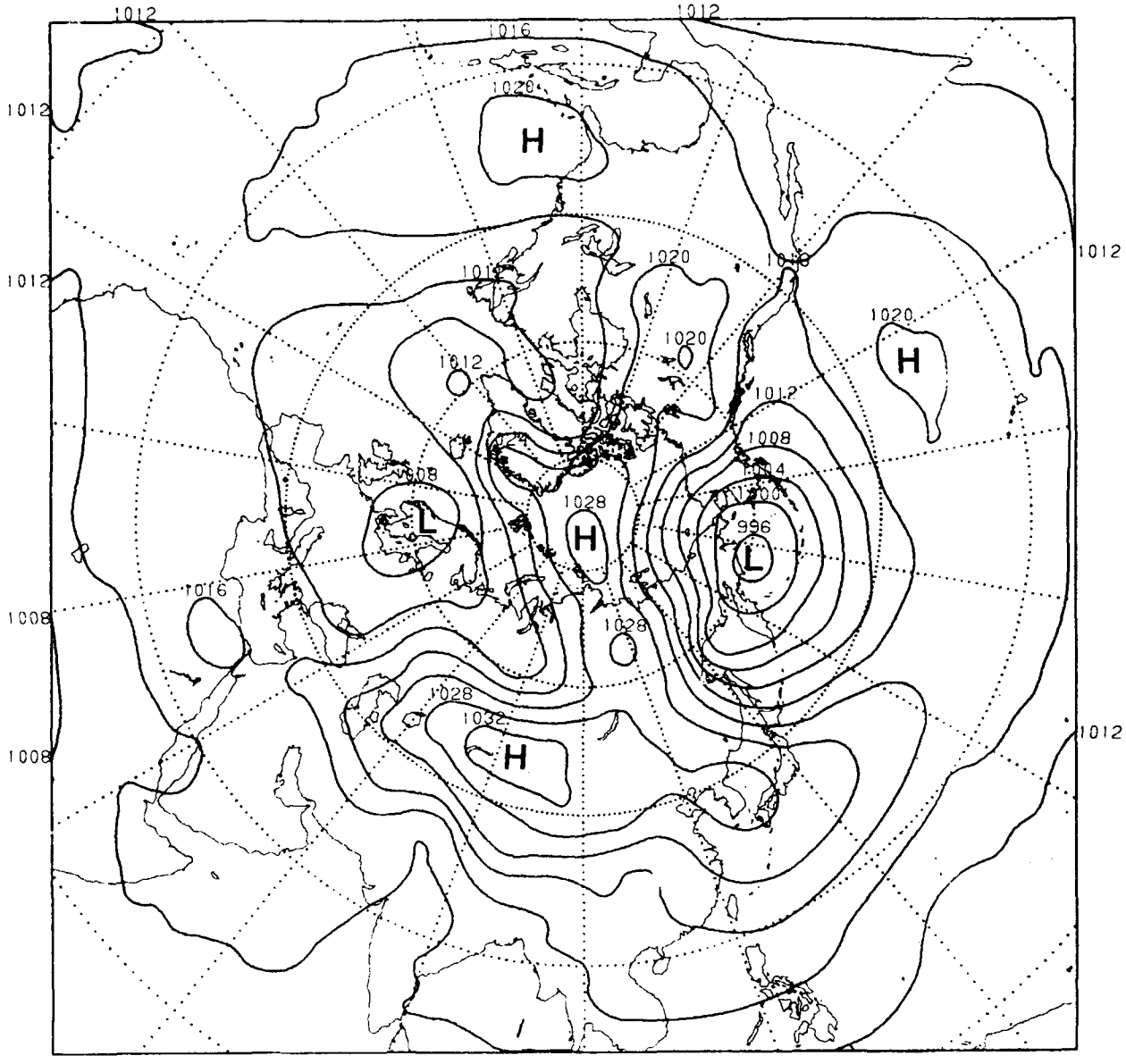

Fig. 11c. Average 48 hour ALLSAT forecast of sea level pressure for 9-31 January 1979 at 1200 GMT.

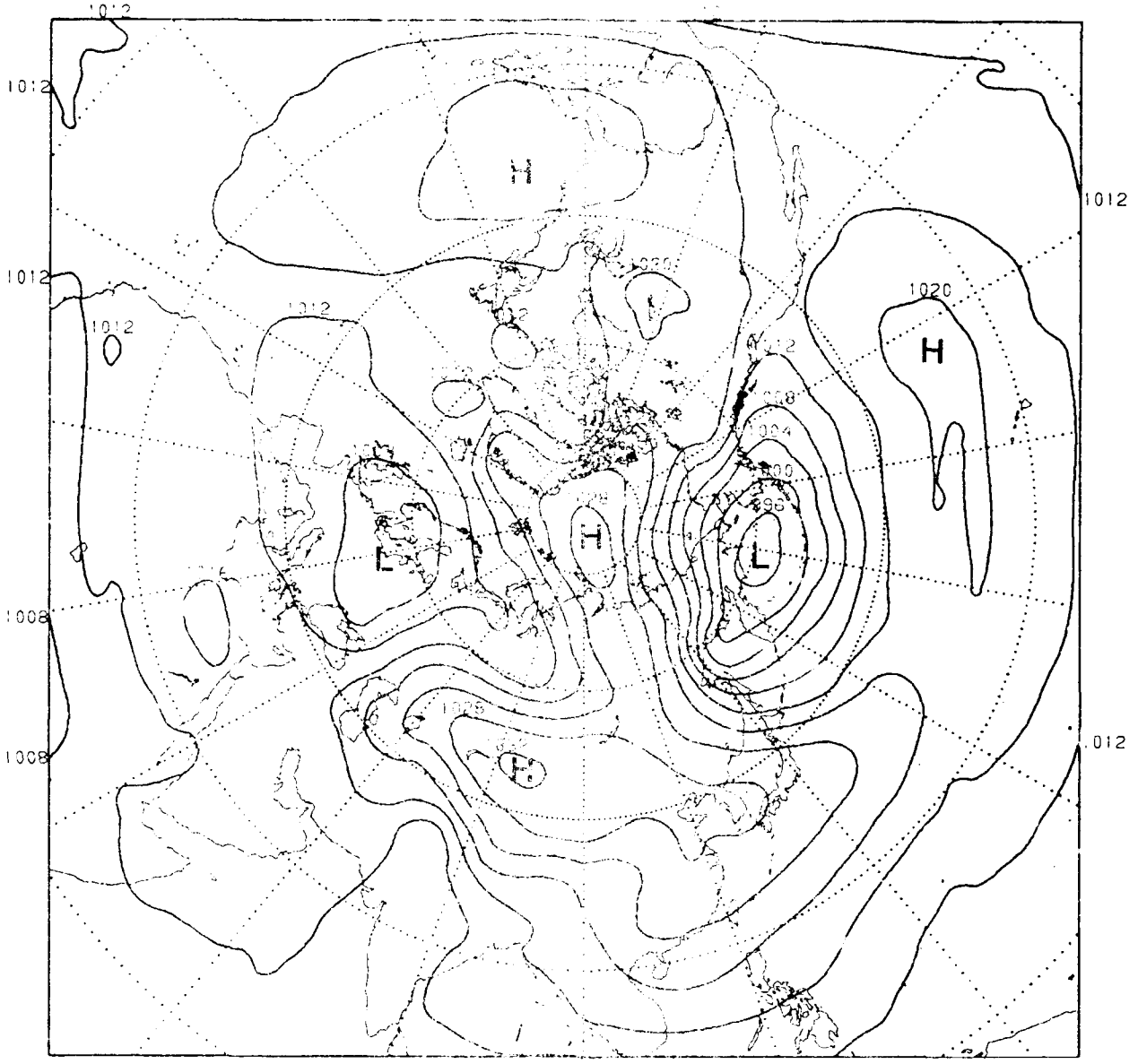

Fig. 11d. Average 48 hour SATCUESS forecast of sea level pressure for 9-31 January 1979 at 1300 GMT.

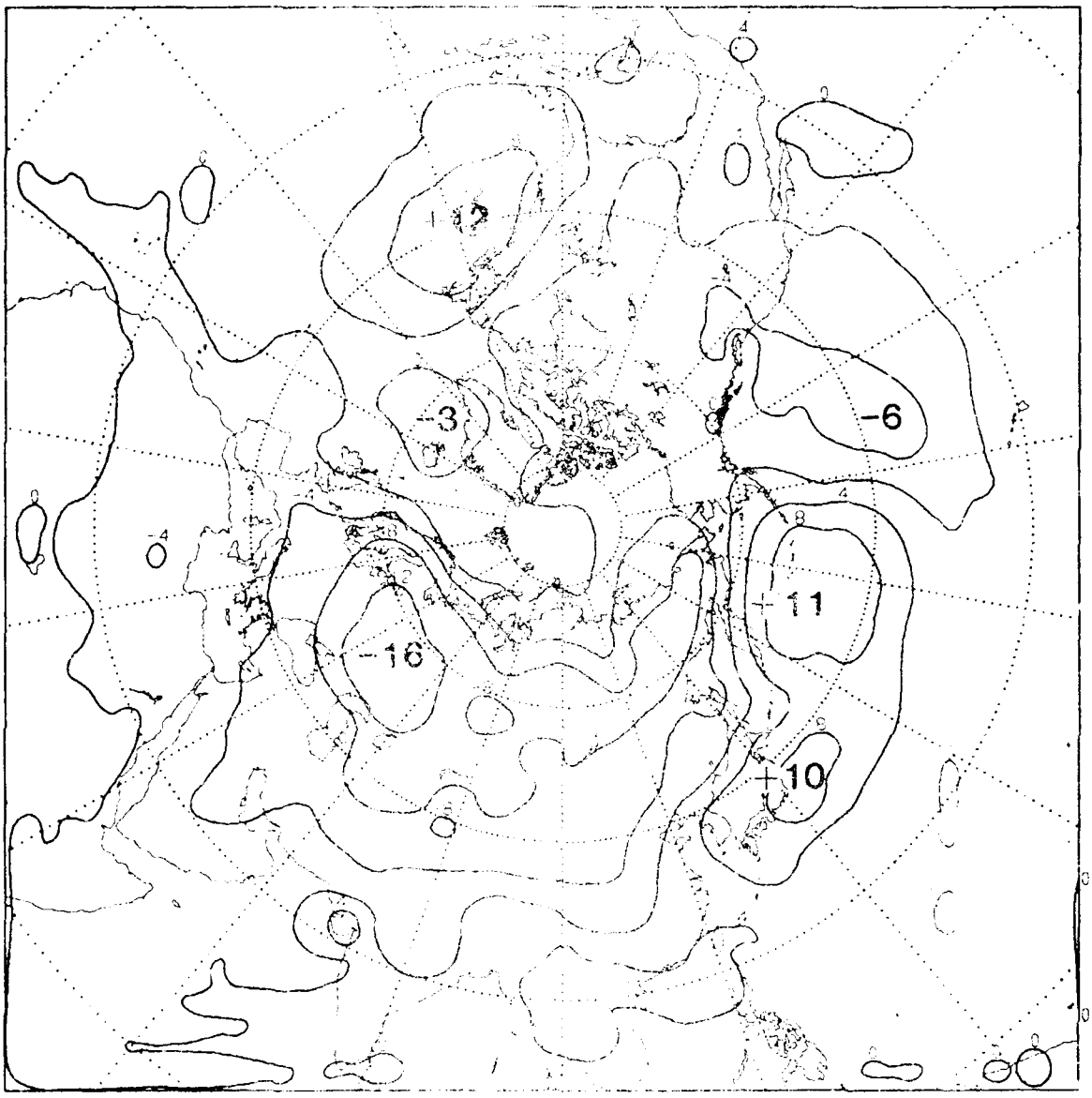

Fig. 12a. Systematic error (mb) of the NOSAT 48 hour forecast of sea level pressure for mid January 1979 at 1200 GMT.

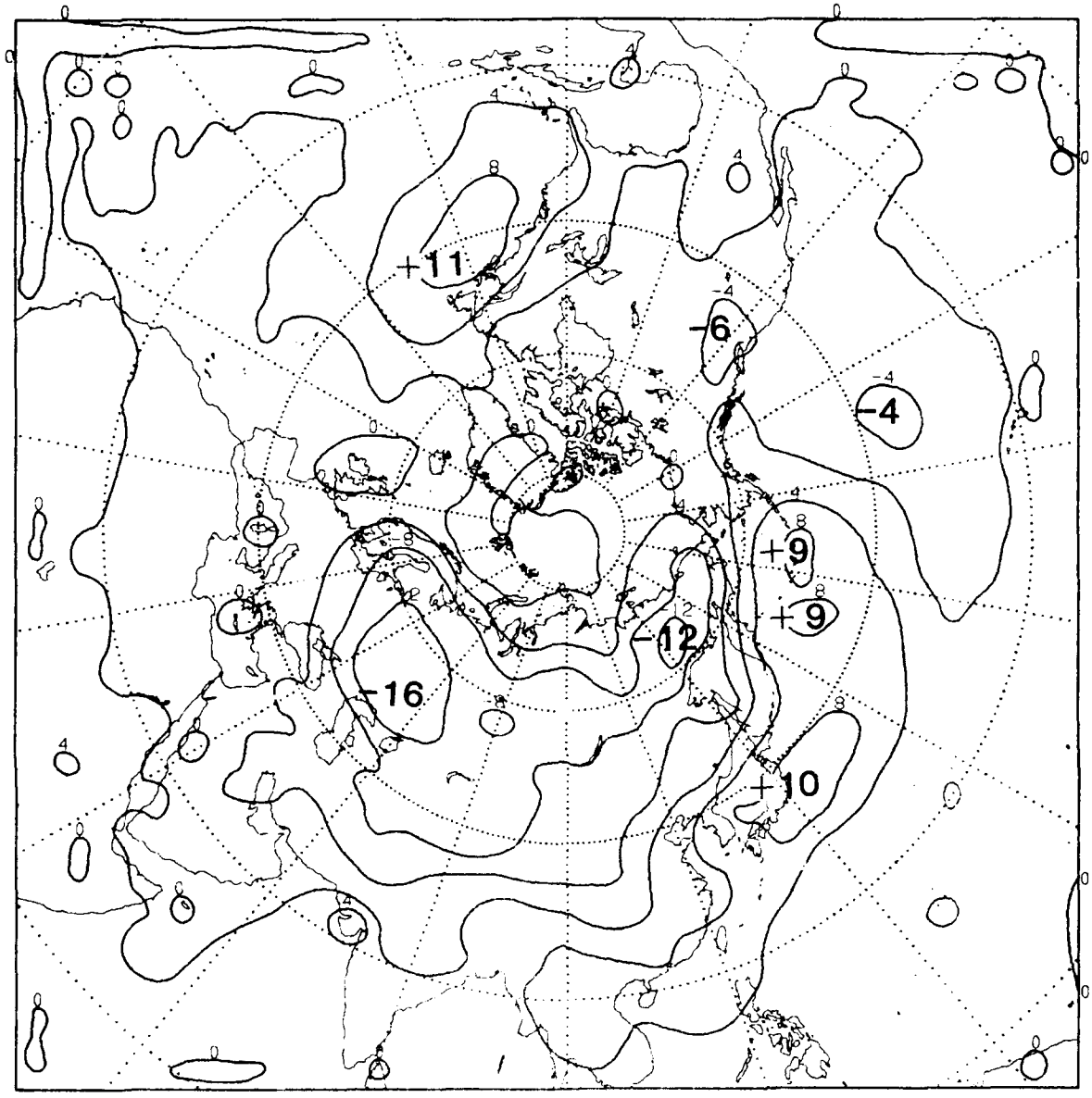

Fig. 12b. Systematic error (mb) of the SAT 48 hour forecast of sea level pressure for 9-31 January 1979 at 1200 GMT.

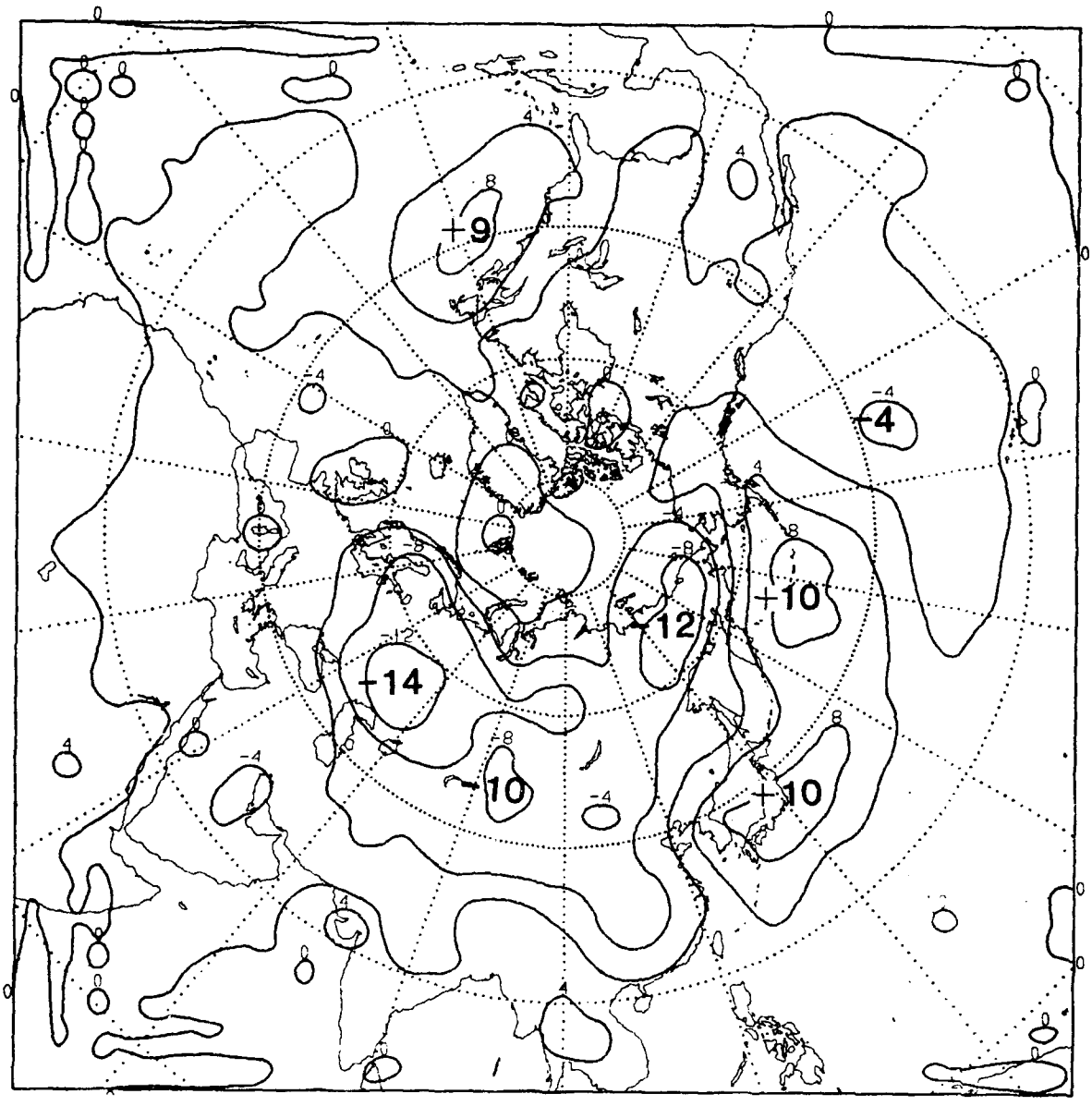

Fig. 12c. Systematic error (mb) of the ALLSAT 48 hour forecast of sea level pressure for 9-31 January 1979 at 1200 GMT.

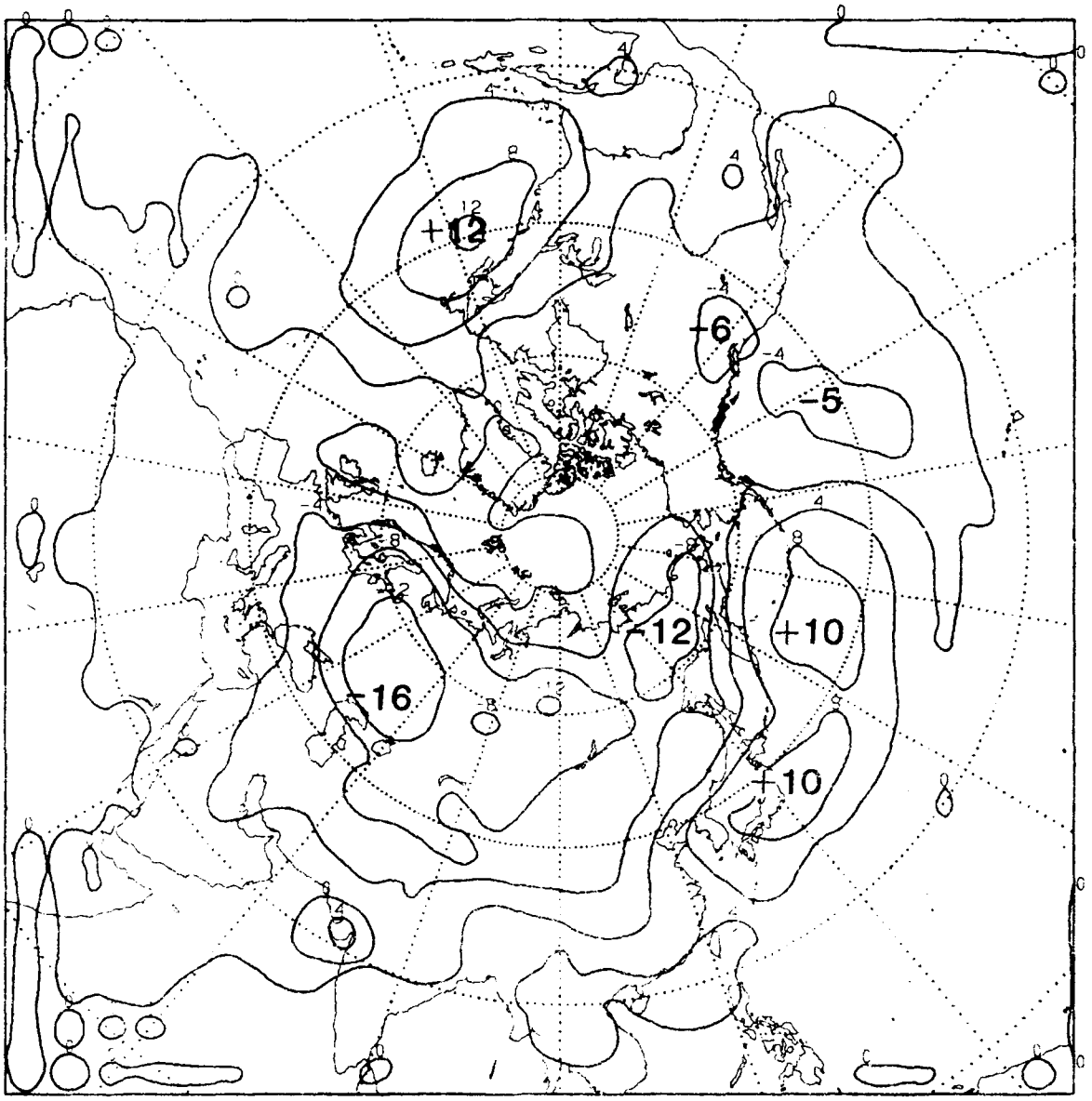

Fig. 12d. Systematic error (mb) of the SATGUESS 48 hour forecast of sea level pressure for 9-31 January 1979 at 1200 GMT.

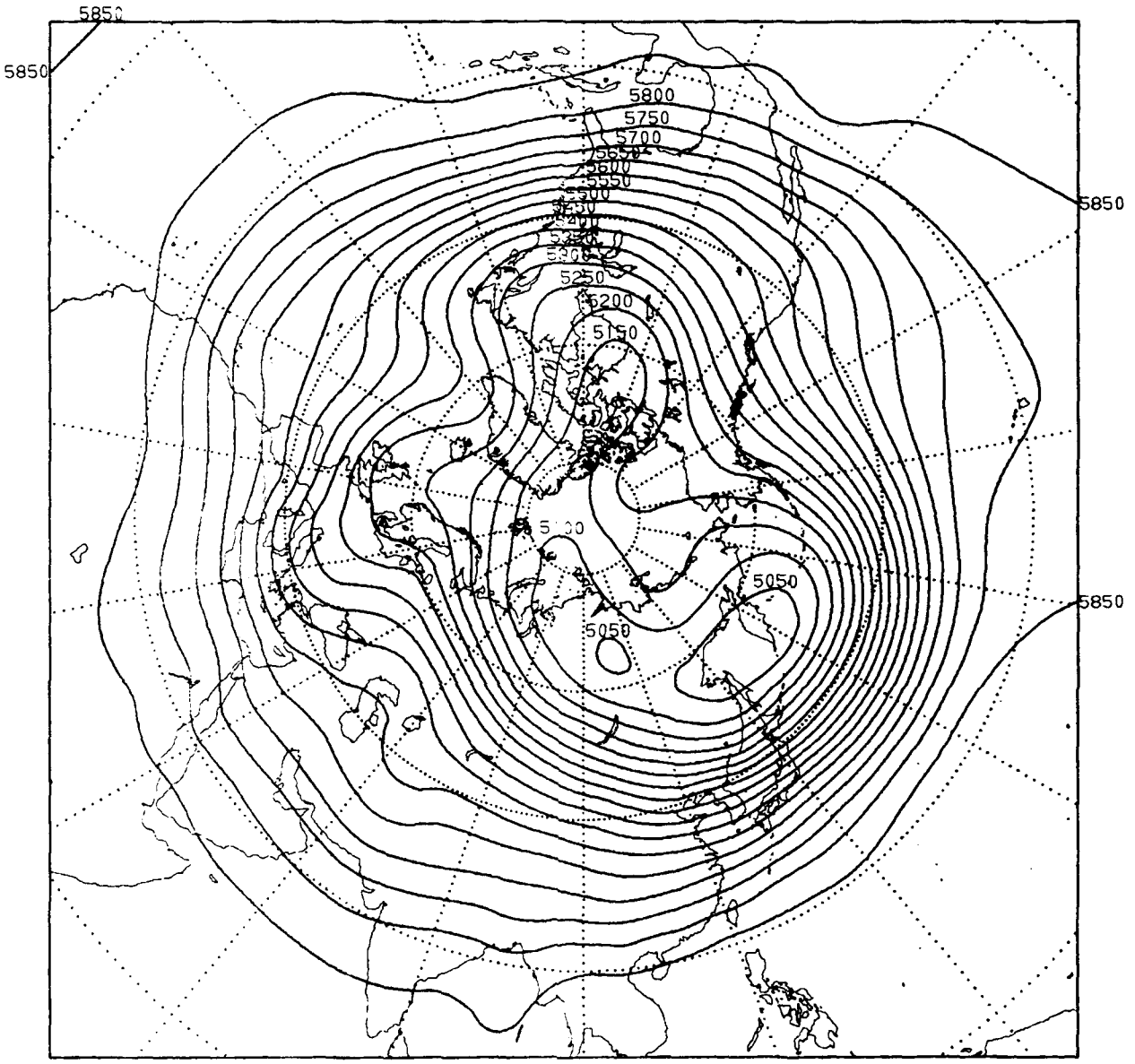

Fig. 13a. Average NOSAT 500 mb height analysis (m) for 9-31 January 1979 at 1200 GMT.

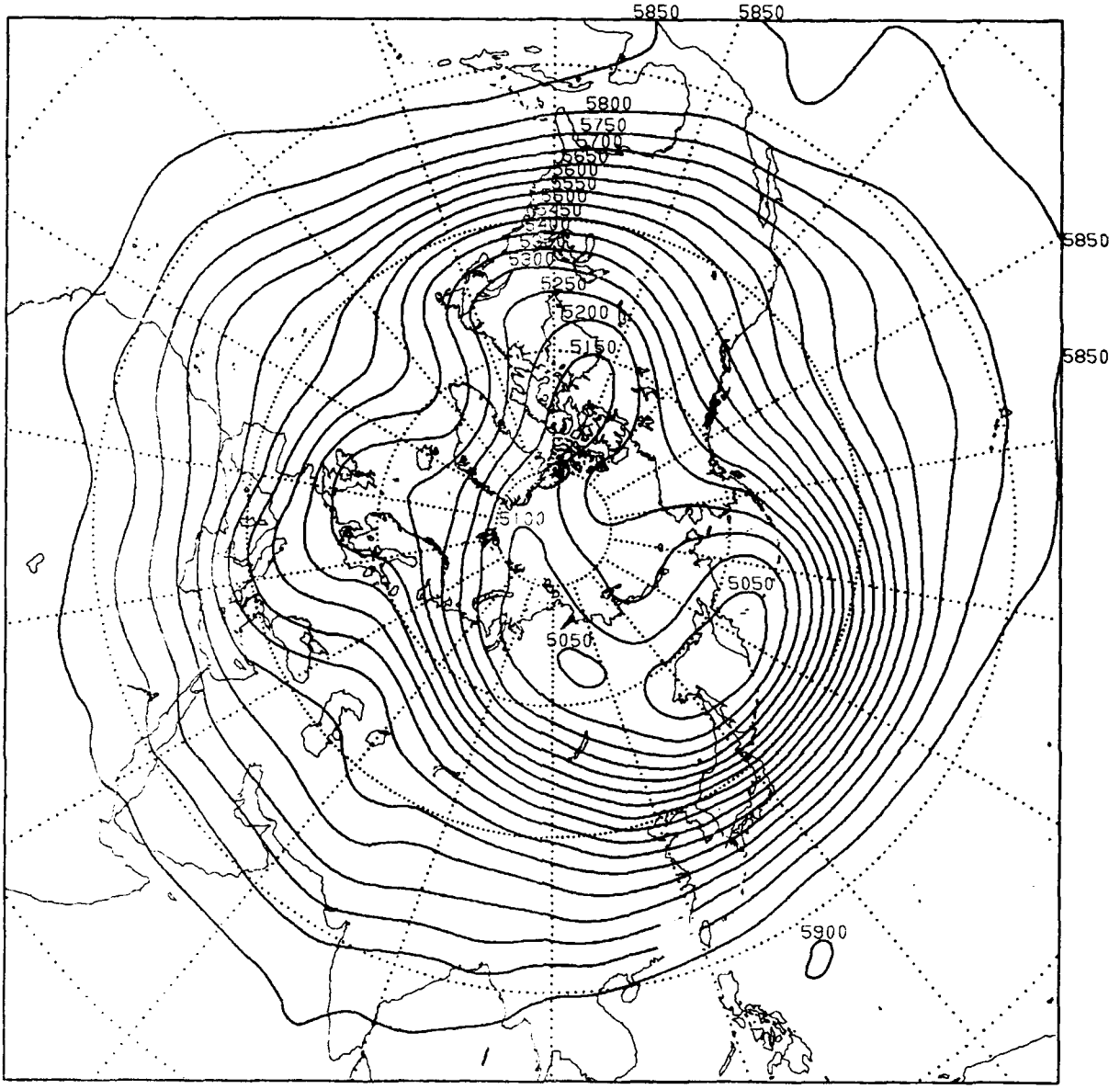

Fig. 13b. Average SAT 500 mb height analysis (m) for 9-31 January 1979 at 1200 GMT.

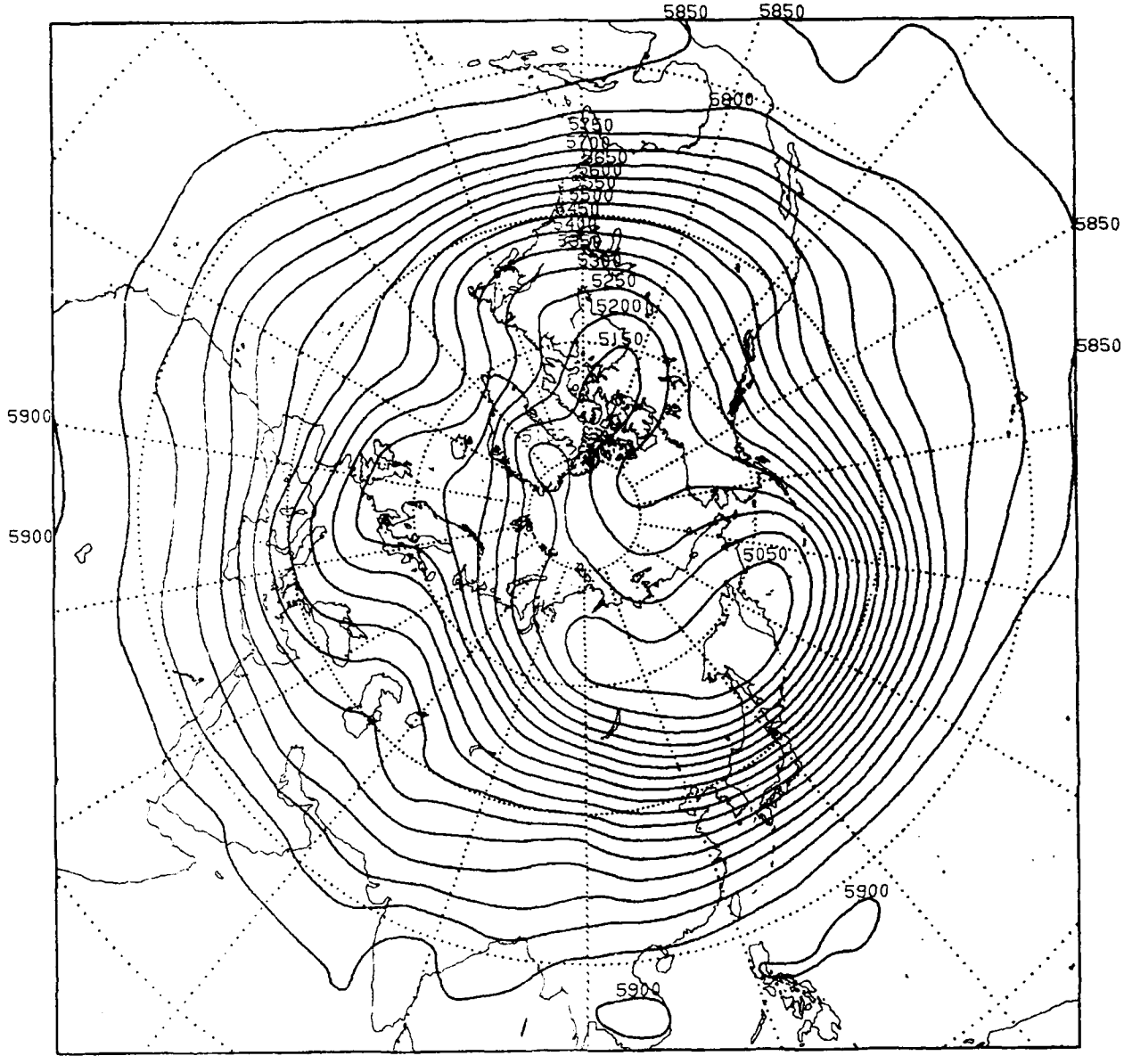

Fig. 13c. Average ALLSAT 500 mb height analysis (m) for 9-31 January 1979 at 1200 GMT.

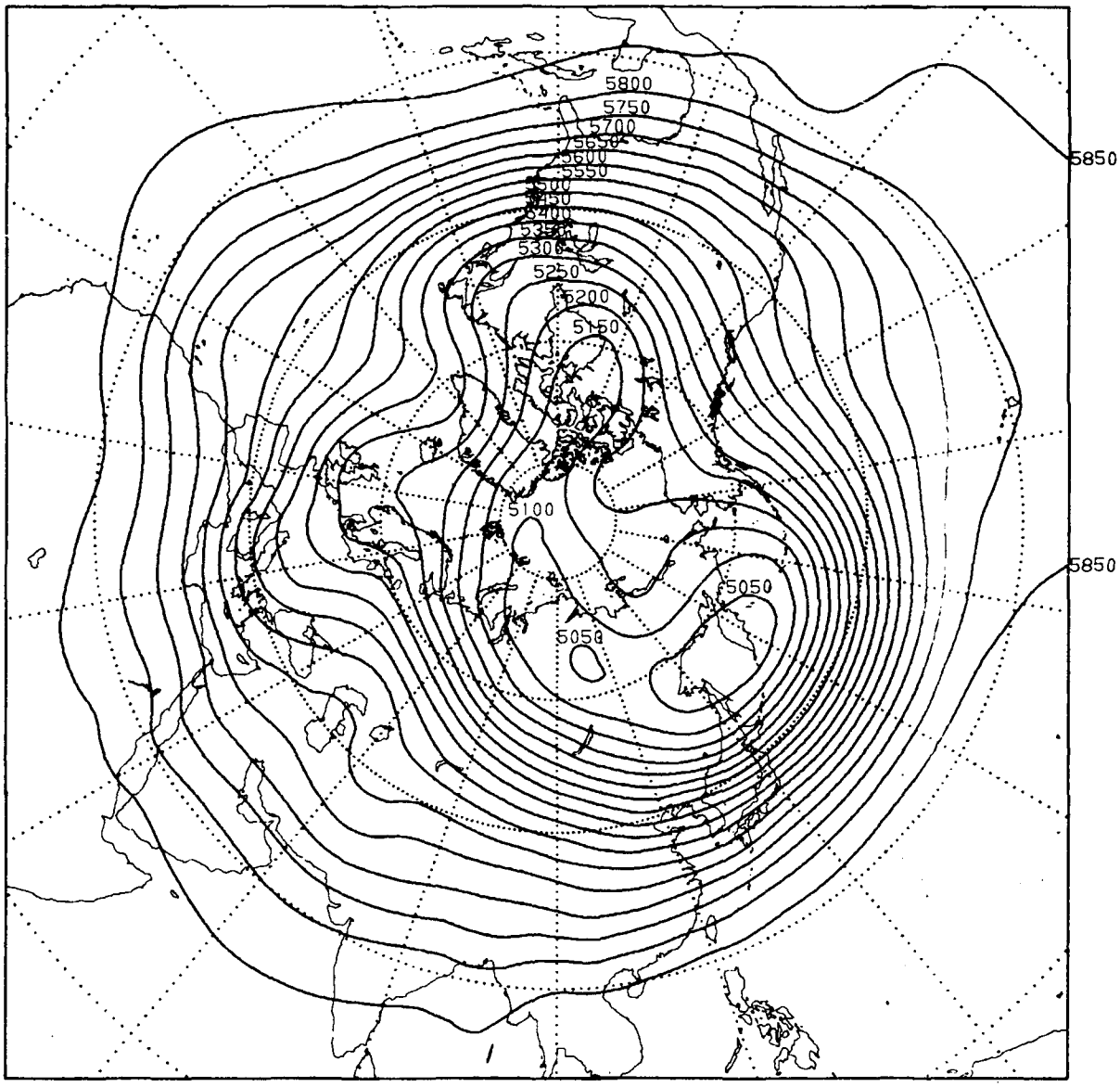

Fig. 13d. Average SATGUESS 500 mb height analysis (m) for 9-31 January 1979 at 1200 GMT.

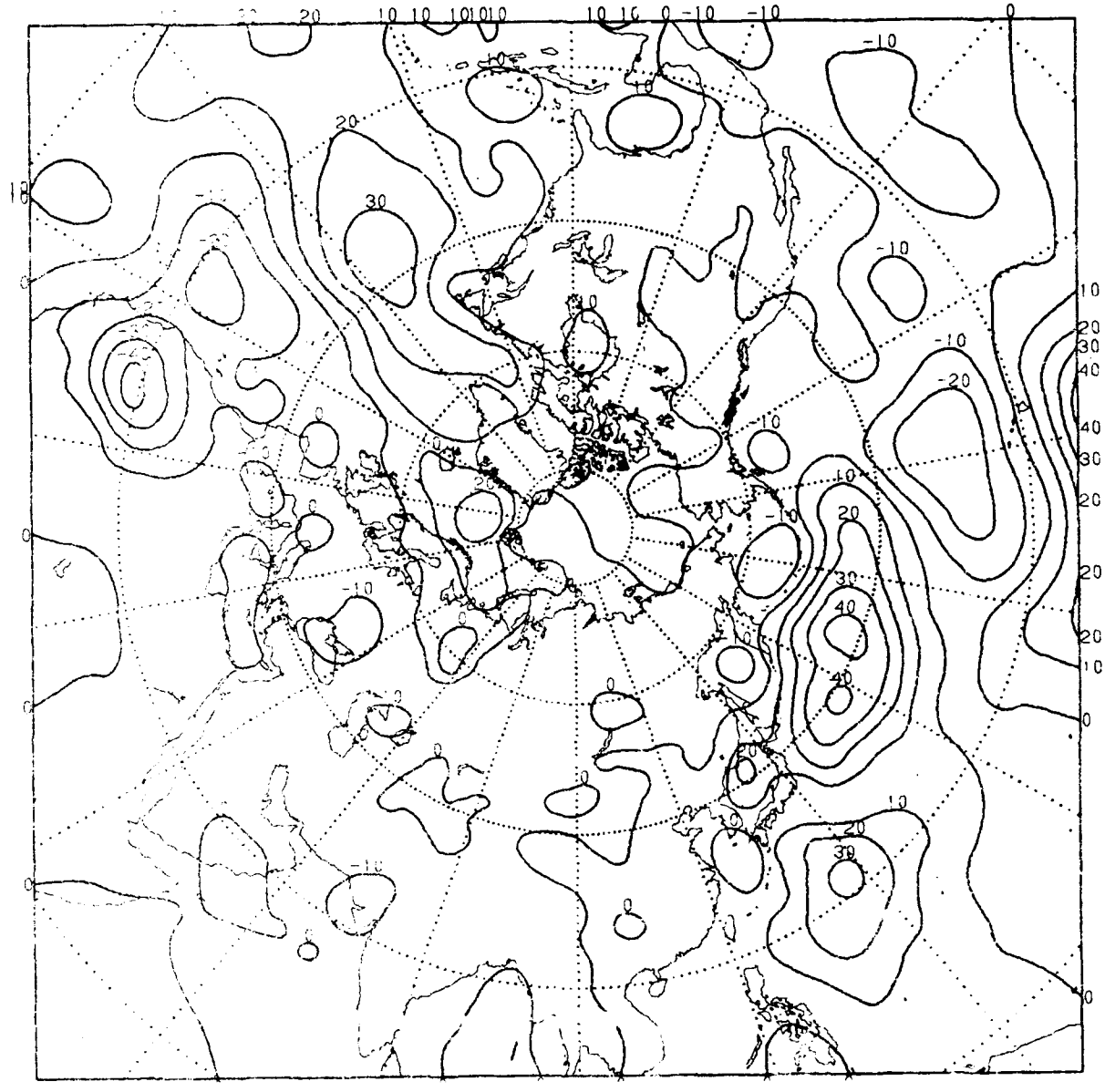

Fig. 14a. Differences between SAT and NOSAT (SAT - NOSAT) average 500 mb analyses (m) for 9-31 January 1979 at 1200 GMT.

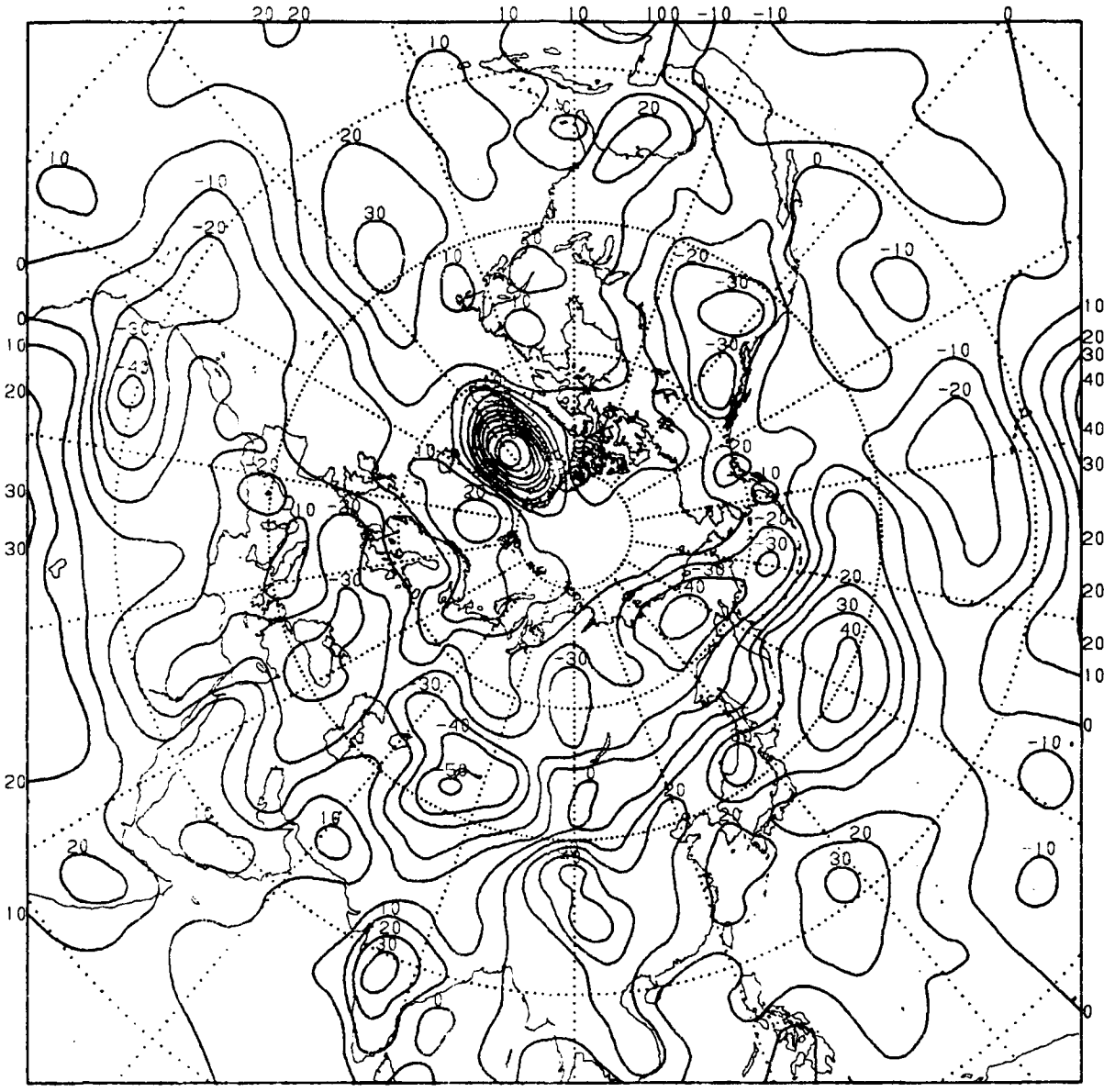

Fig. 14b. Differences between ALLSAT and NOSAT (ALLSAT - NOSAT) average 500 mb analyses (m) for 9-31 January 1979 at 1200 GMT.

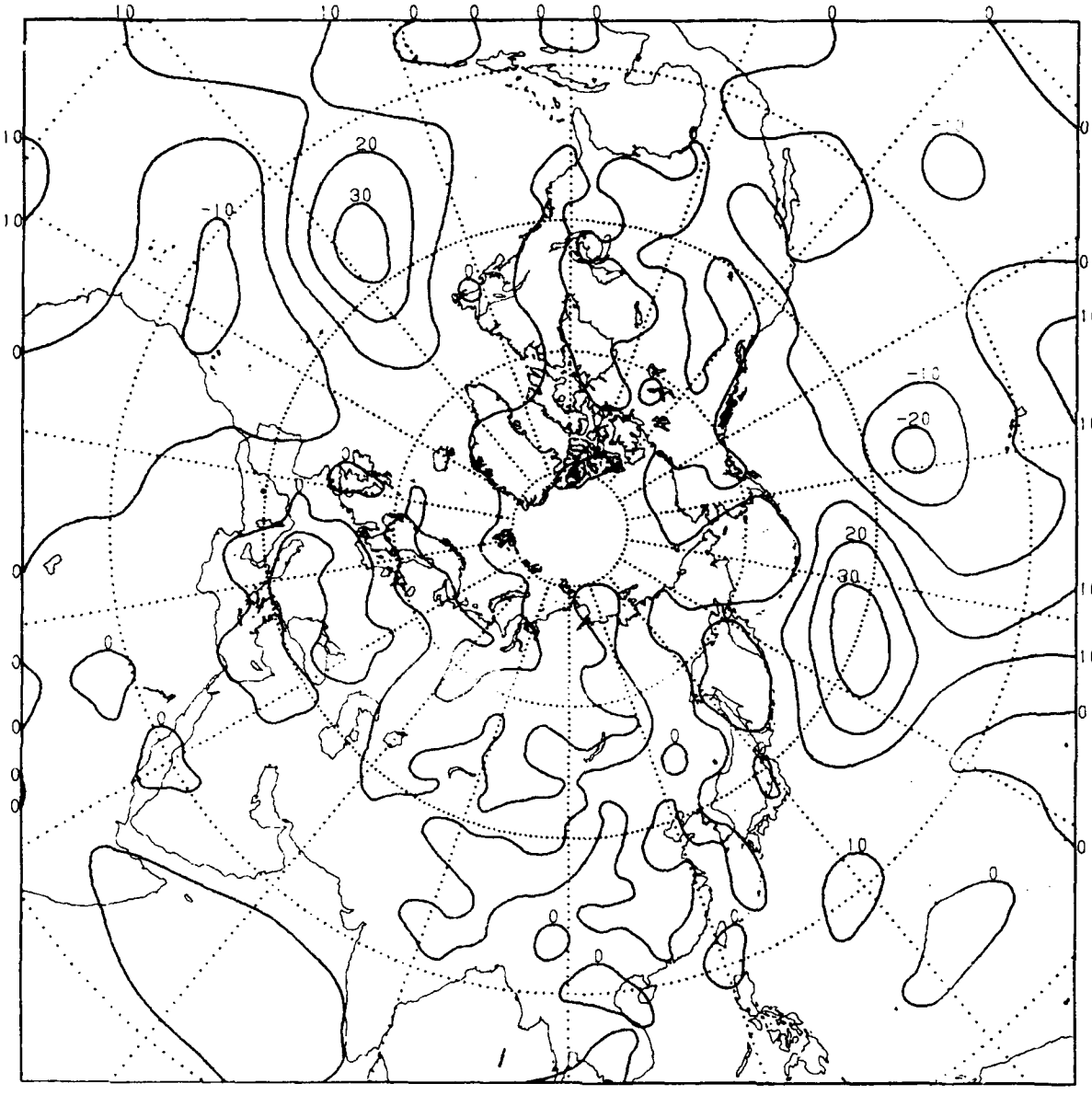

Fig. 14c. Differences between SATGUESS and NOSAT (SATGUESS - NOSAT) average 500 mb analyses (m) for 9-31 January 1979 at 1200 GMT.

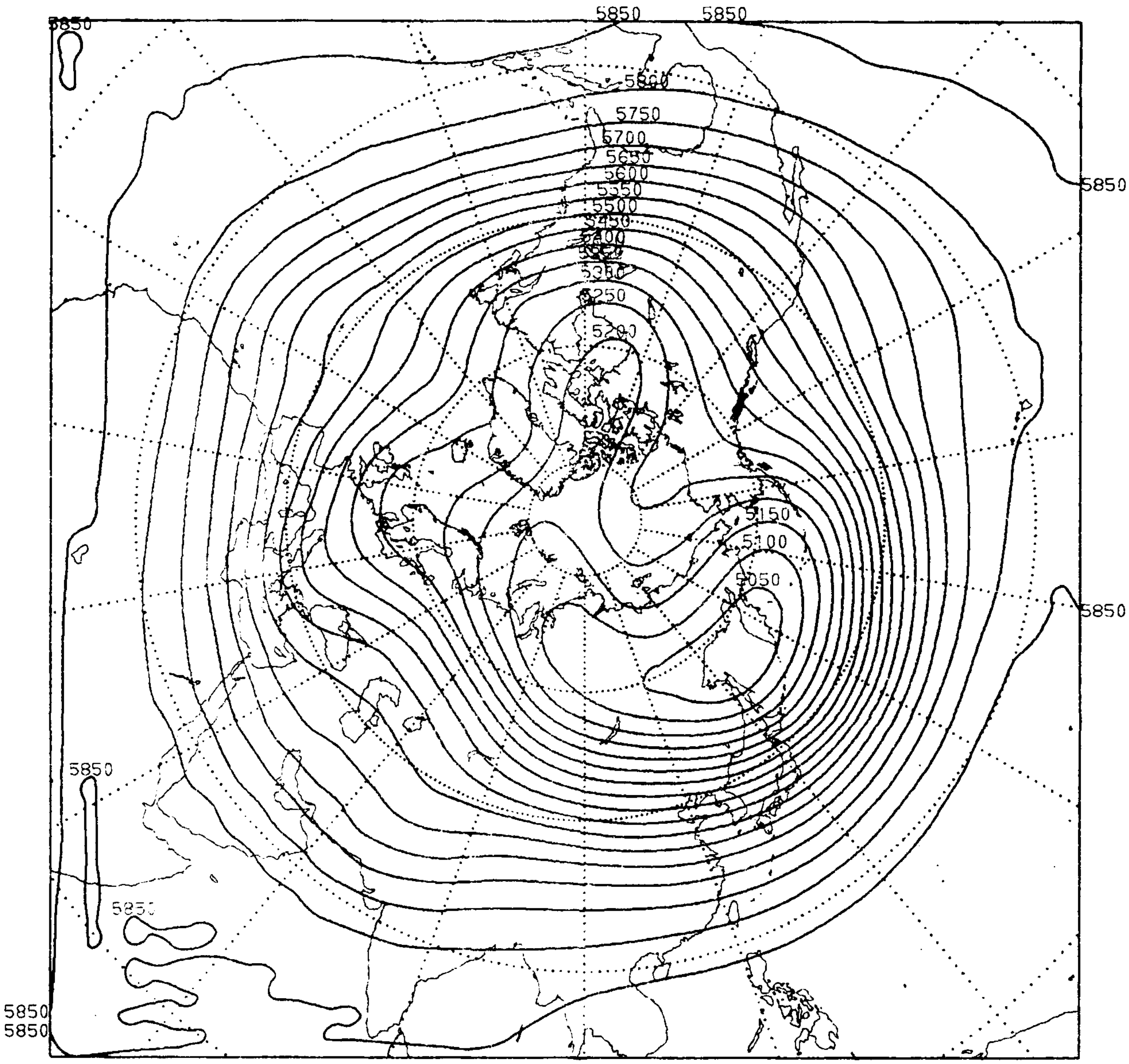

Fig. 15a. Average 48 hour NOSAT forecast of 500 mb height (m) for 9-31 January 1979 at 1200 GMT.

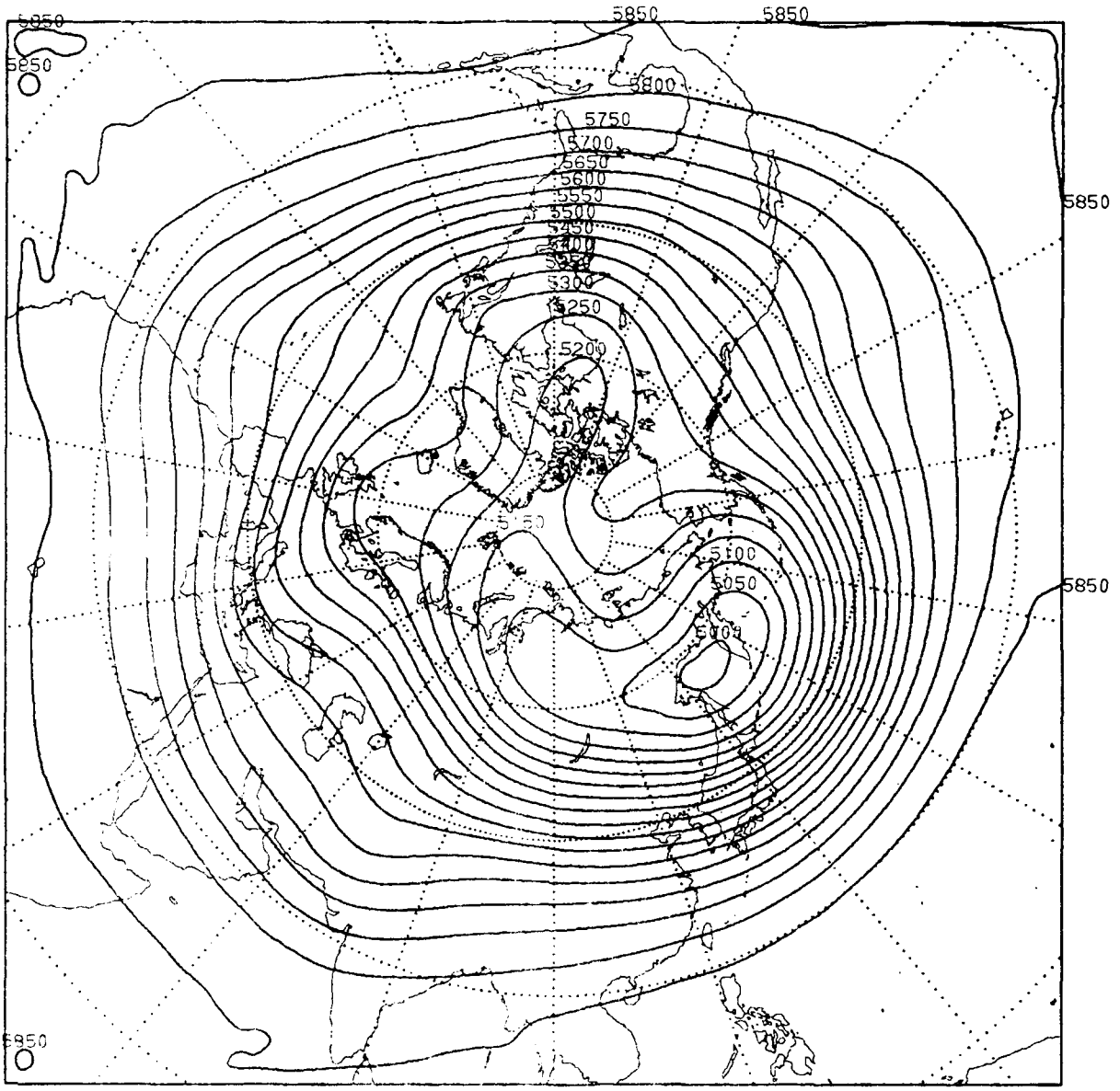

Fig. 15b. Average 48 hour SAT forecast of 500 mb height (m) for 9-31 January 1979 at 1200 GMT.

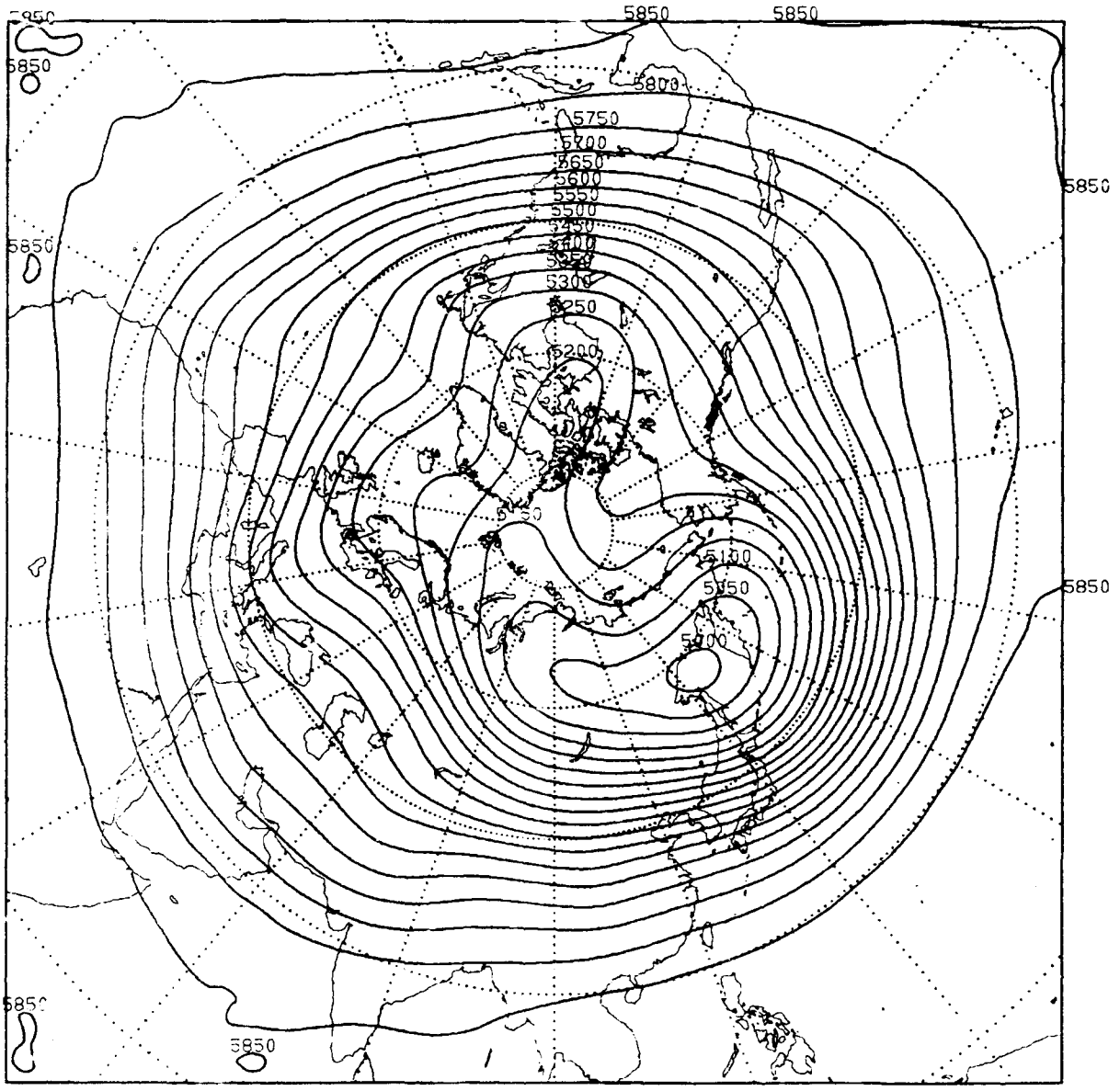

Fig. 15c. Average 48 hour ALLSAT forecast of 500 mb height (m) for 9-31 January 1979 at 1200 GMT.

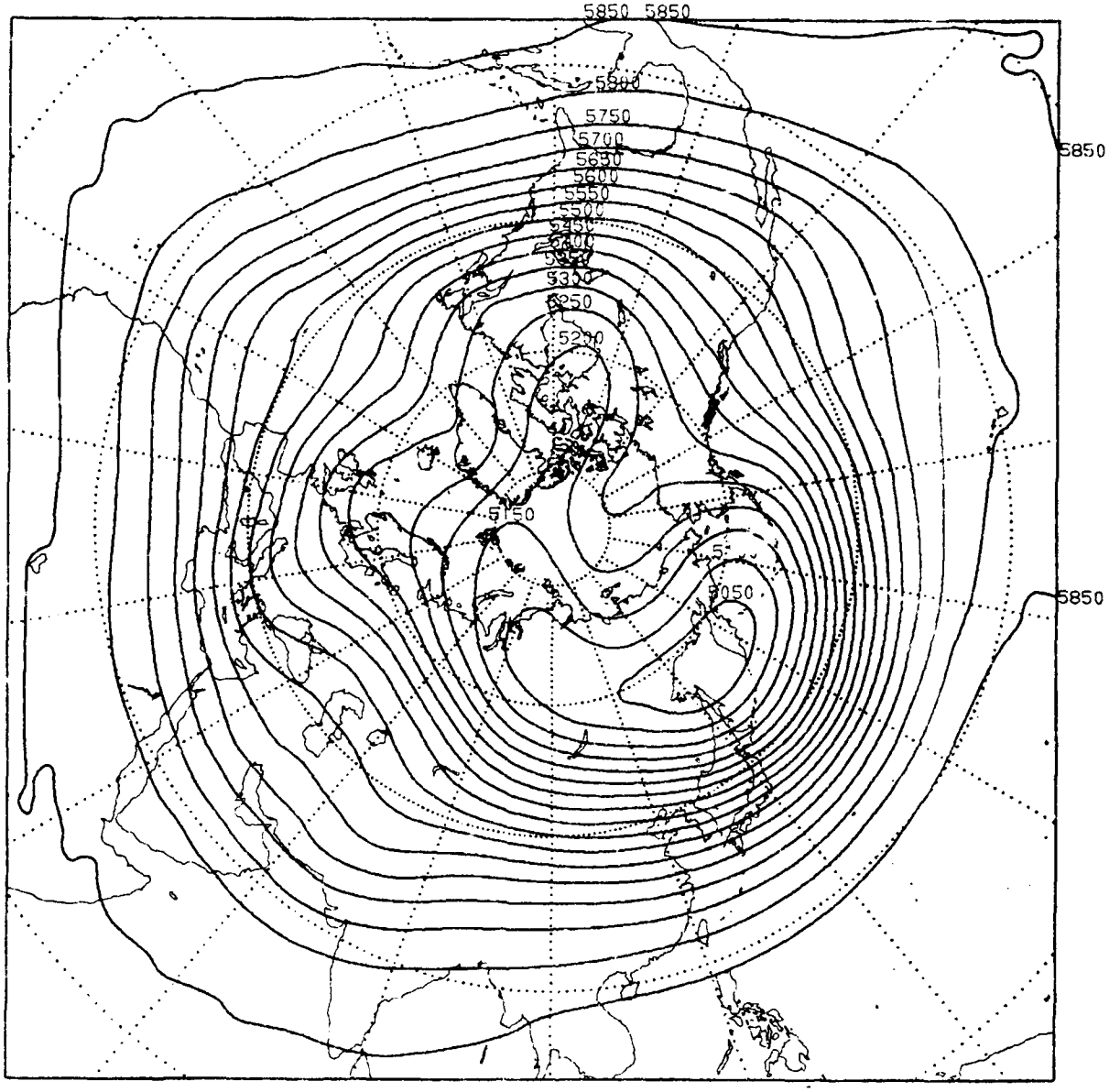

Fig. 15d. Average 48 hour SATGUESS forecast of 500 mb height (m) for 9-31 January 1979 at 1200 GMT.

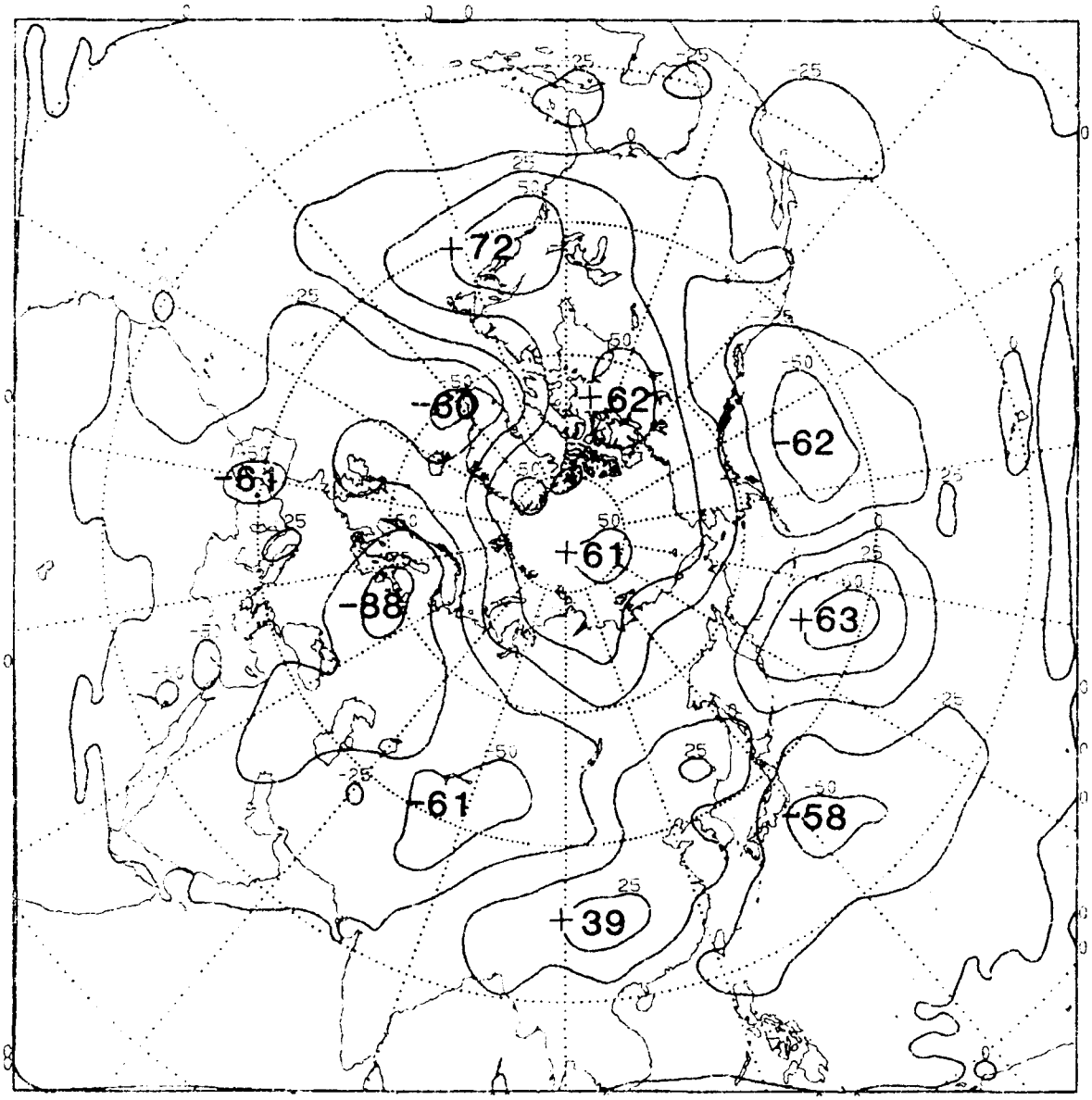

Fig. 16a. Systematic error (m) of the NOSAT 48 hour forecast of 500 mb height for 9-31 January 1979 at 1200 GMT.

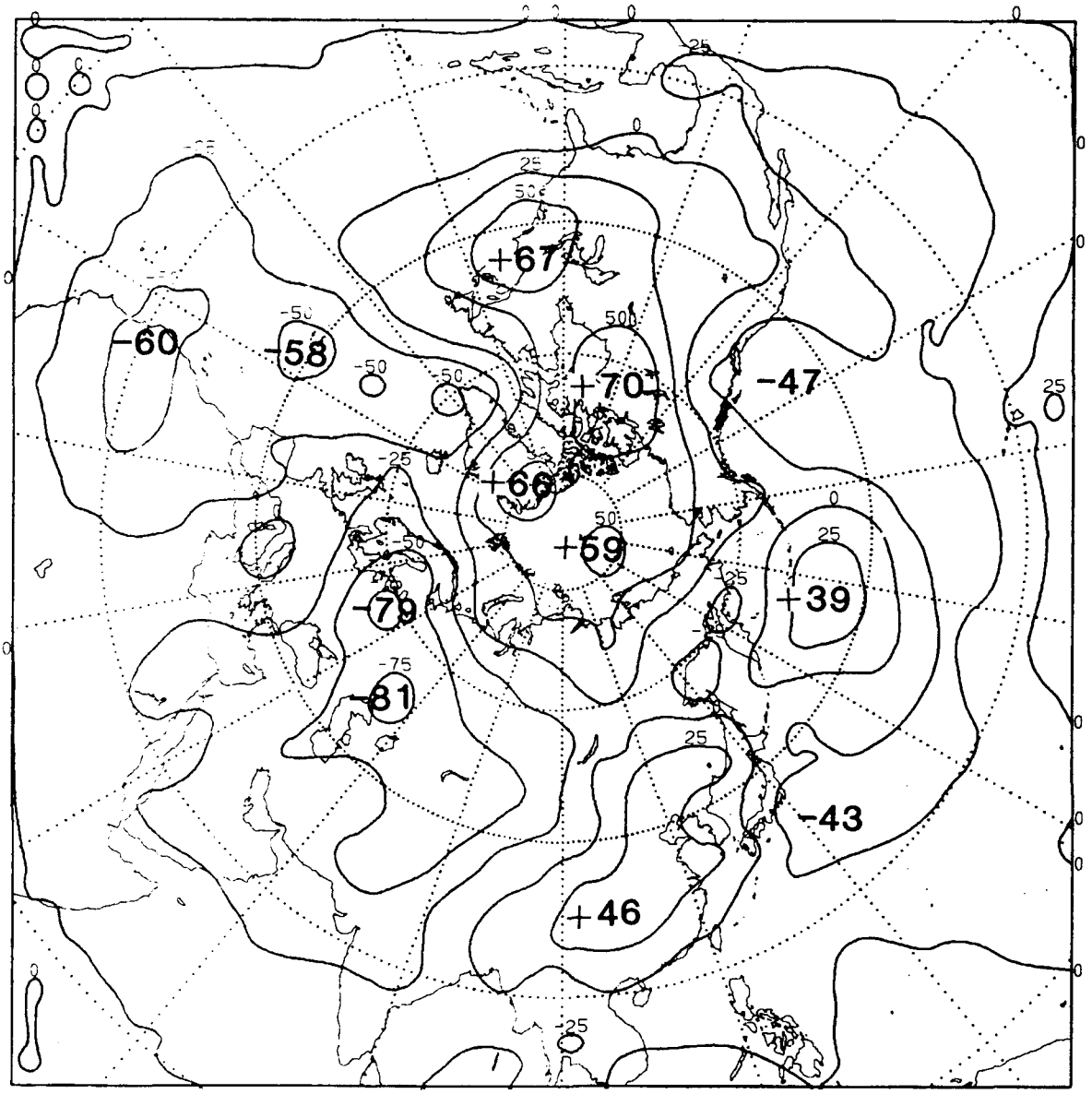

Fig. 16b. Systematic error (m) of the SAT 48 hour forecast of 500 mb height for 9-31 January 1979 at 1200 GMT.

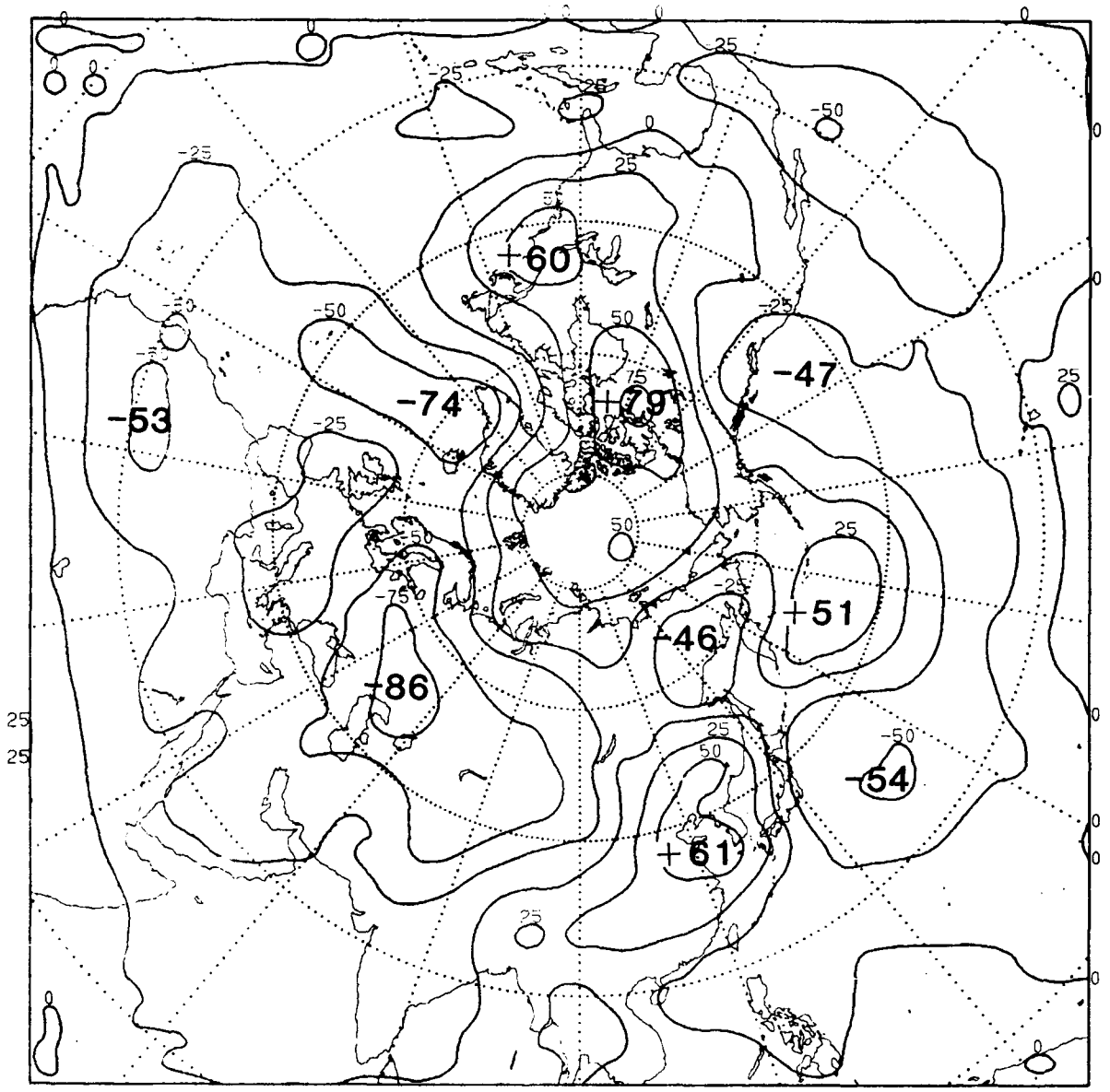

Fig. 16c. Systematic error (m) of the ALLSAT 48 hour forecast of 500 mb height for 9-31 January 1979 at 1200 GMT.

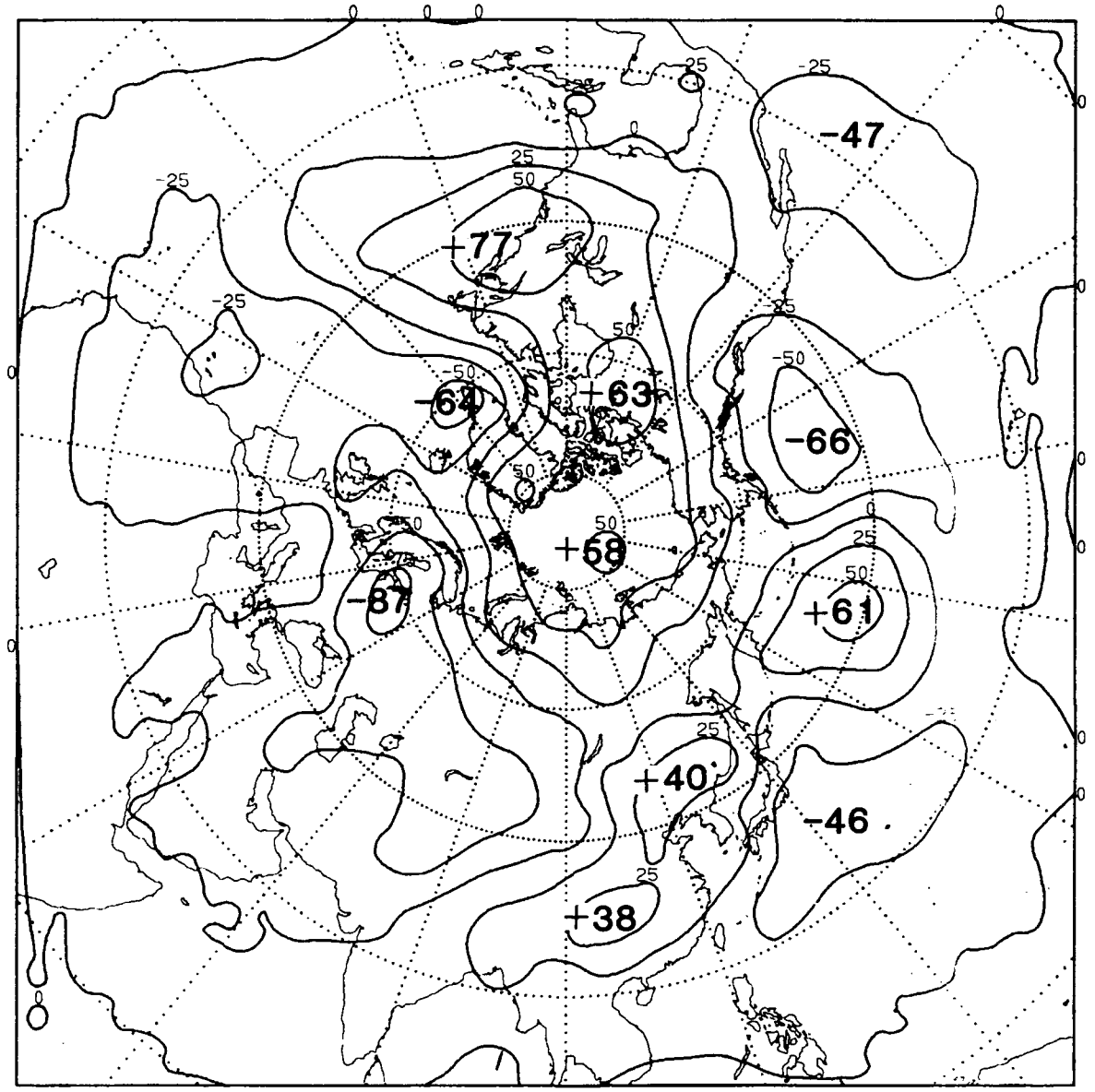

Fig. 16d. Systematic error (m) of the SATGUESS 48 hour forecast of 500 mb height for 9-31 January 1979 at 1200 GMT.

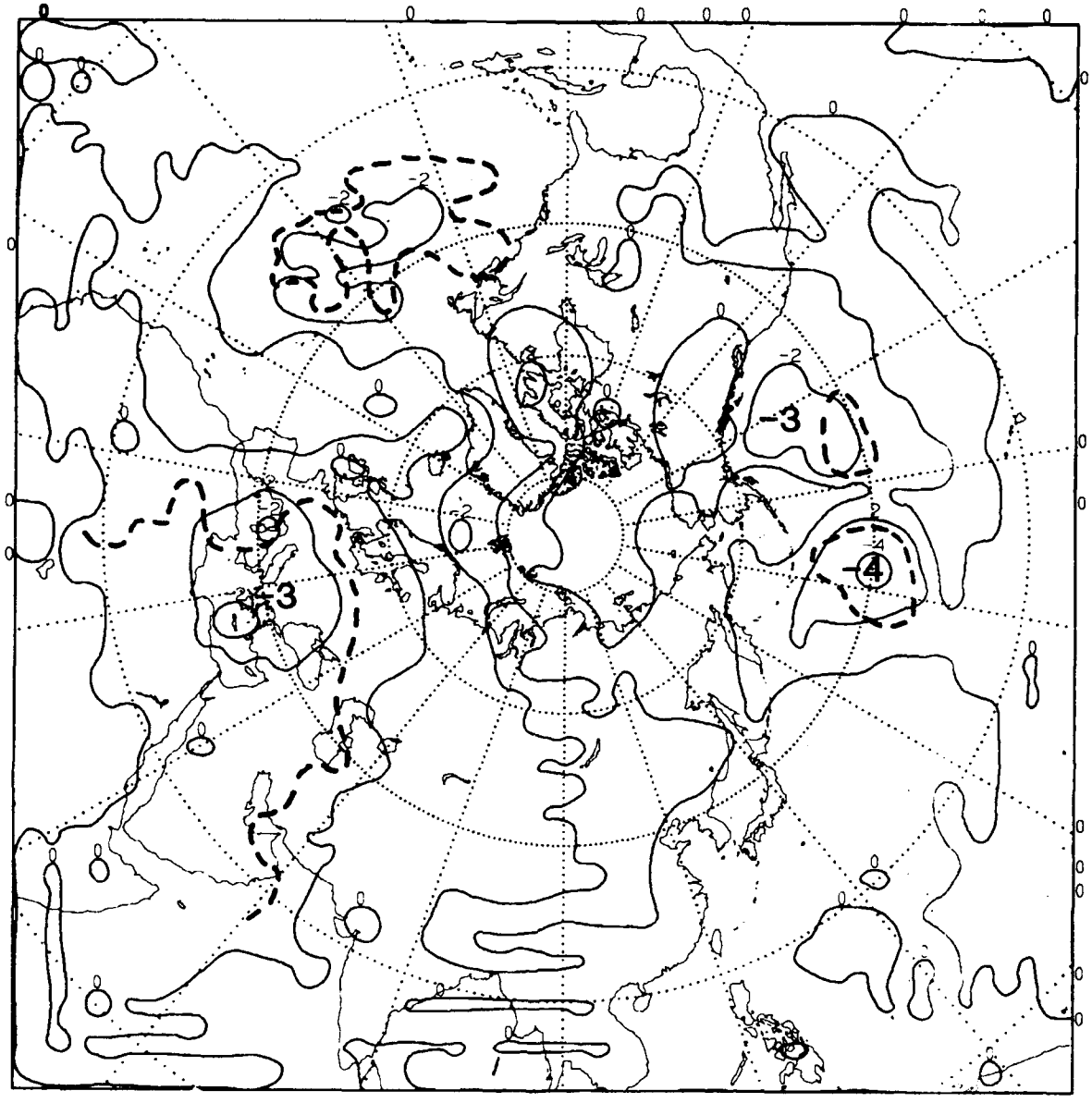

Fig. 17a. Differences between absolute values of SAT and NOSAT systematic errors (|SAT| - |NOSAT|) of 48 hour forecasts of sea level pressure for 9-31 January 1979 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

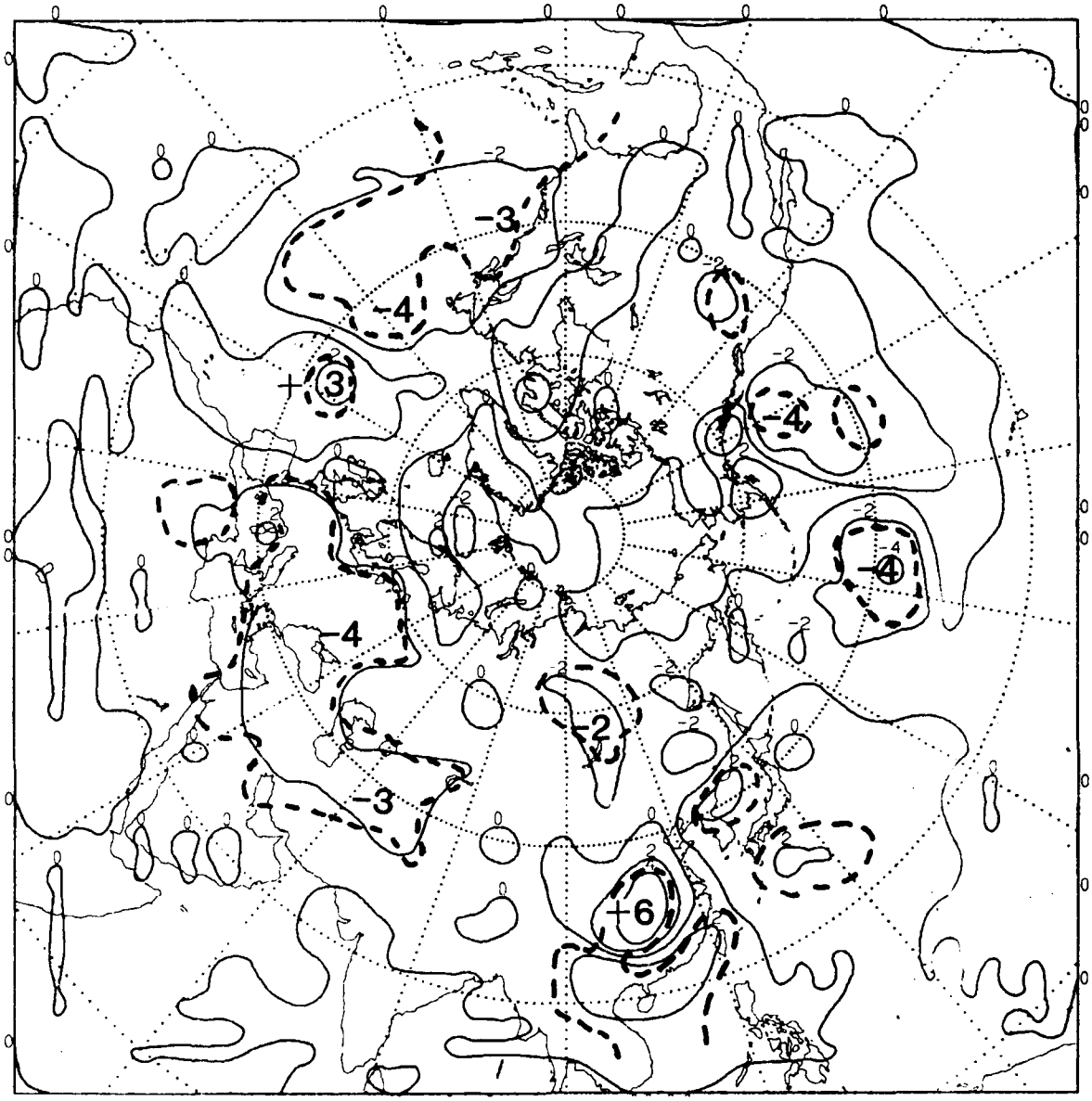

Fig. 17b. Differences between absolute values of ALLSAT and NOSAT systematic errors (|ALLSAT|-|NOSAT|) of 48 hour forecasts of sea level pressure for 9-31 January 1979 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

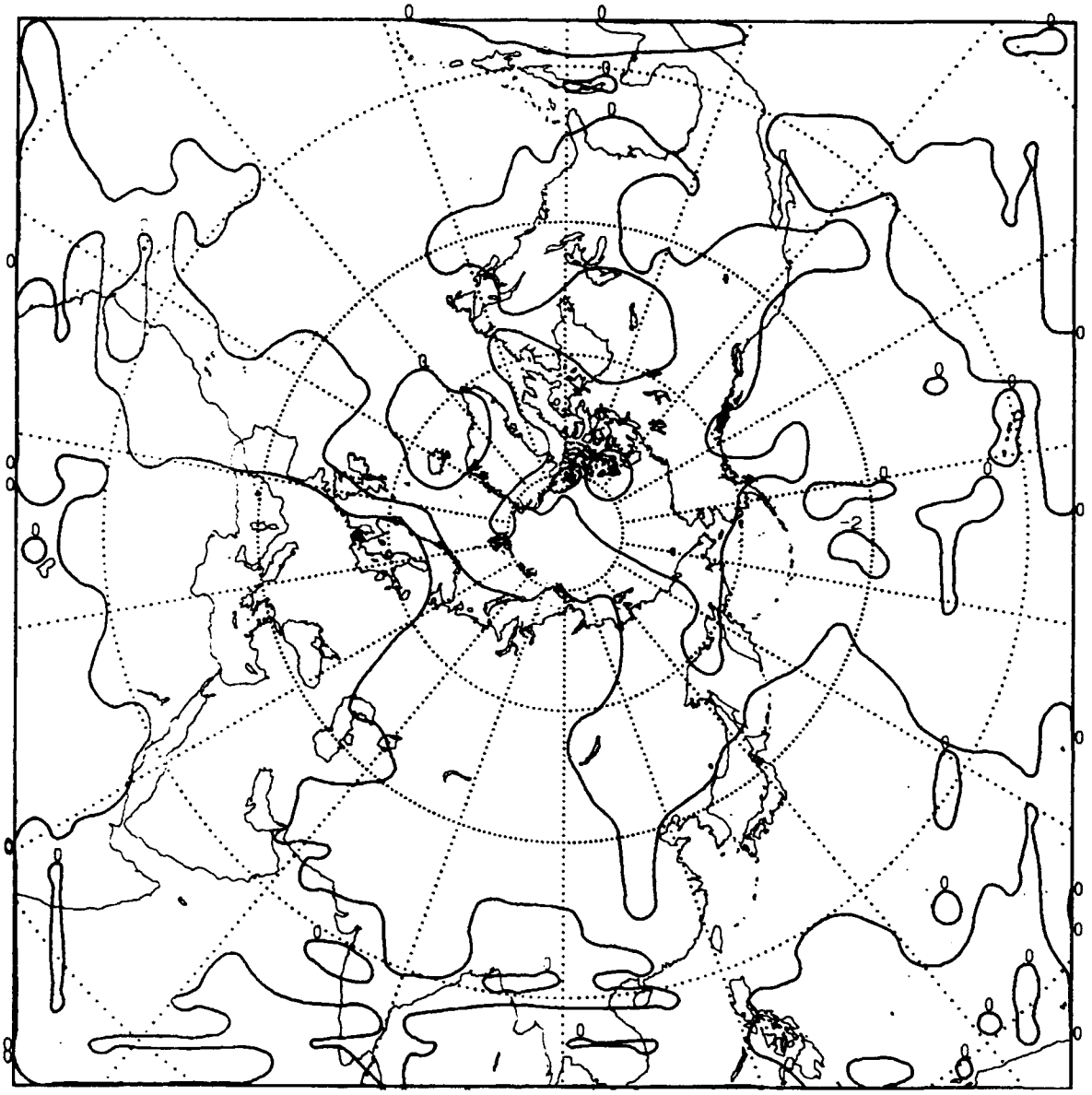

Fig. 17c. Differences between absolute values of SATGUESS and NOSAT systematic errors (|SATGUESS| - |NOSAT|) of 48 hour forecasts of sea level pressure for 9-31 January 1979 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

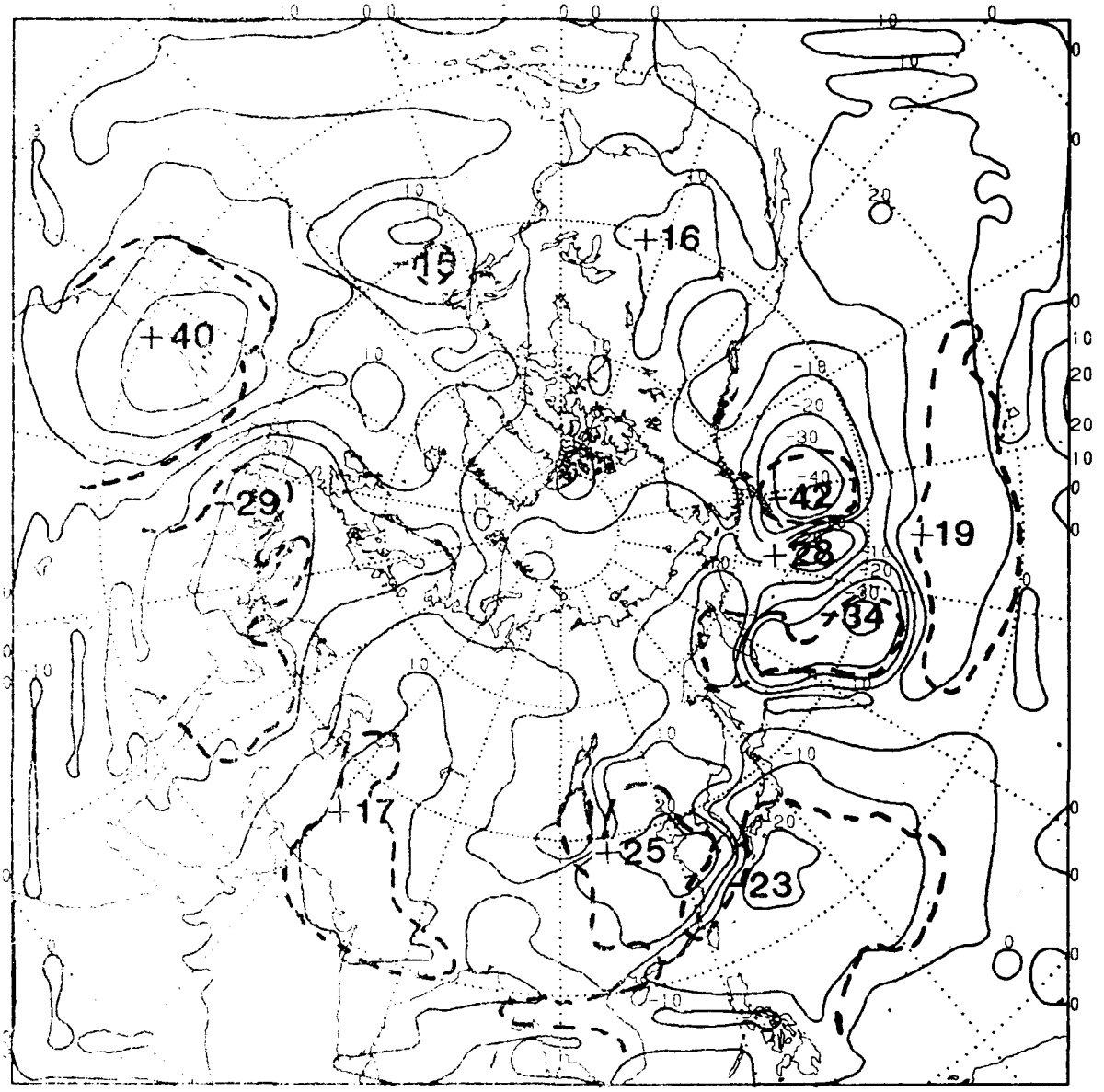

Fig. 18a. Differences between absolute values of SAT and NOSAT systematic errors (|SAT| - |NOSAT|) of 48 hour forecasts of 500 mb height for 9-31 January 1979 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

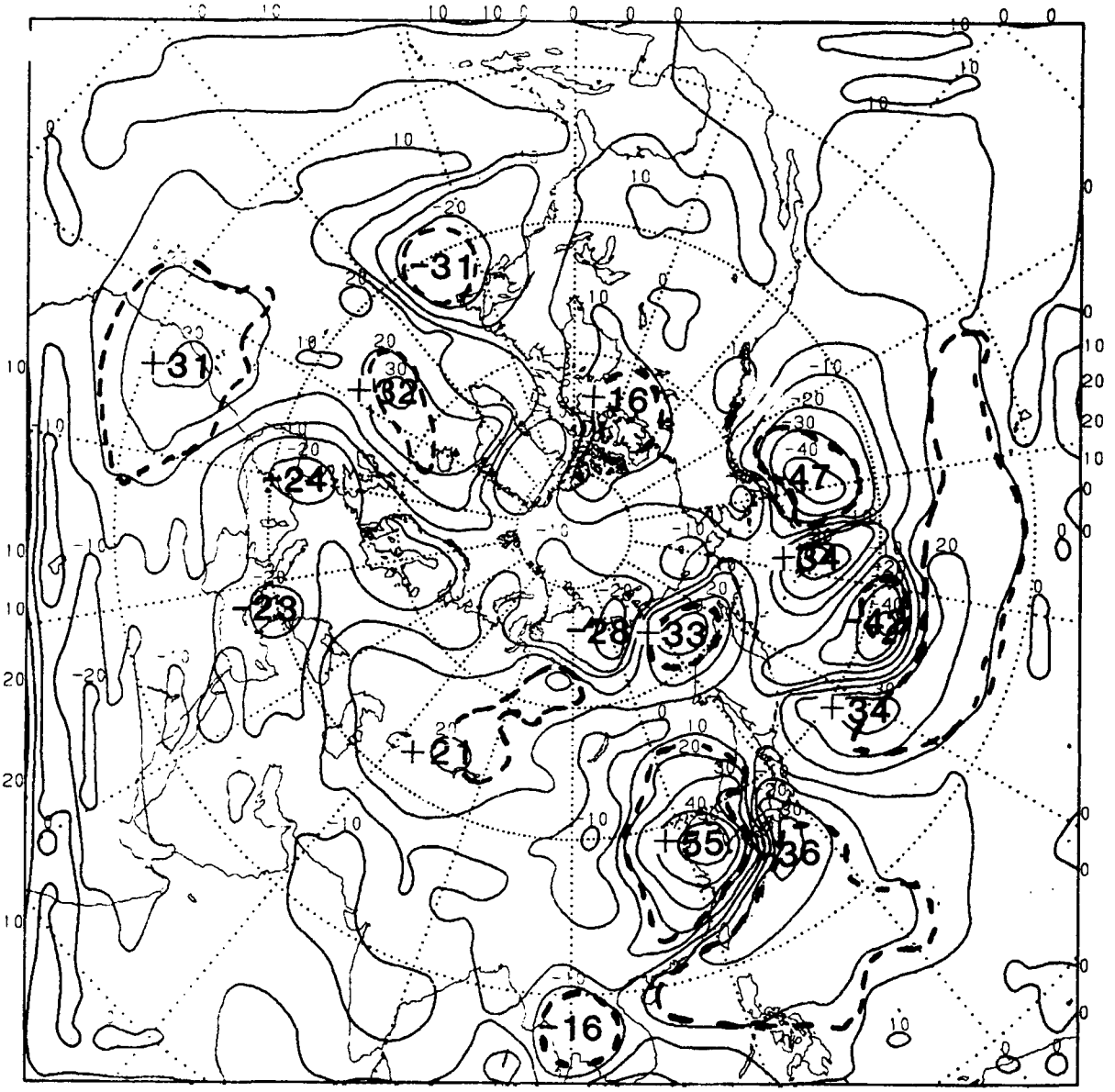

Fig. 18b. Differences between absolute values of ALLSAT and NOSAT systematic errors (|ALLSAT| - |NOSAT|) of 48 hour forecasts of 500 mb height for 9-31 January 1979 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

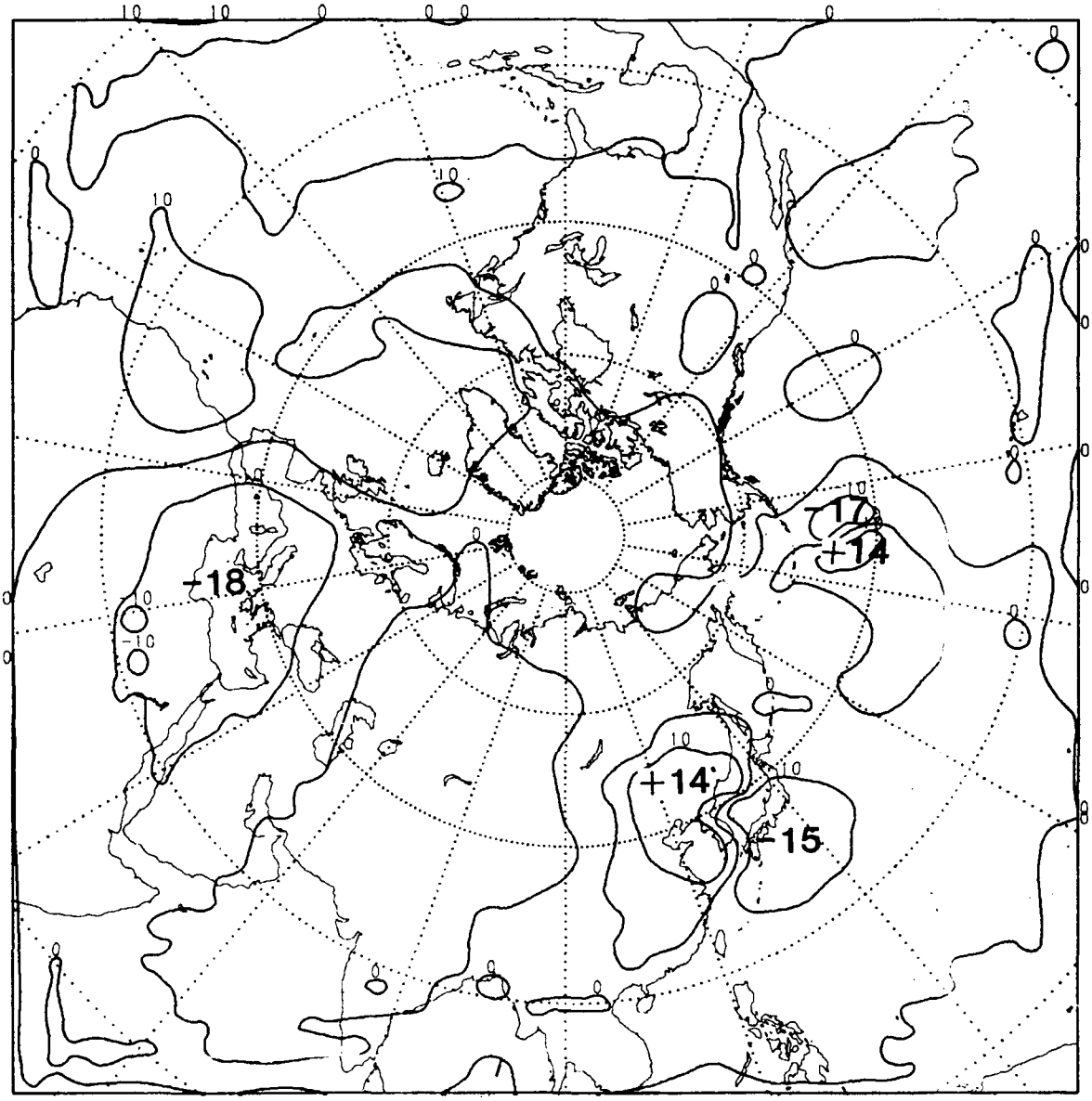

Fig. 18c. Differences between absolute values of SATGUESS and NOSAT systematic errors (|SATGUESS| - |NOSAT|) of 48 hour forecasts of 500 mb height for 9-31 January 1979 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

Contents

Table

Figures

level pressure for 16-28 December 1987 at 1200 GMT.

15b. Systematic error (mb) of the SAT 48 hour forecast of sea level pressure for 16-28 December 1987 at 1200 GMT.

16a. Average NOSAT 500 mb height analysis (m) for 16-28 December 1987 at 1200 GMT.

16b. Average SAT 500 mb height analysis (m) for 16-28 December 1987 at 1200 GMT.

17. Differences between SAT and NOSAT (SAT - NOSAT) average 500 mb analyses (m) for 16-28 December 1987 at 1200 GMT.

18a. Average 48 hour NOSAT forecast of 500 mb height (m) for 16-28 December 1987 at 1200 GMT.

18b. Average 48 hour SAT forecast of 500 mb height (m) for 16-28 December 1987 at 1200 GMT.

19a. Systematic error (m) of the NOSAT 48 hour forecast of 500 mb height for 16-28 December 1987 at 1200 GMT.

19b. Systematic error (m) of the SAT 48 hour forecast of 500 mb height for 16-28 December 1987 at 1200 GMT.

20a. Differences between absolute values of SAT and NOSAT systematic errors ($|SAT| - |NOSAT|$) of 48 hour forecasts of sea level pressure for 16-28 December 1987 at 1200 GMT.

20b. Differences between absolute values of SAT and NOSAT systematic errors ($|SAT| - |NOSAT|$) of 48 hour forecasts of 500 mb height (m) for 16-28 December 1987 at 1200 GMT.

Part II - December 1987 Data

1. Introduction

This part of the report describes the impact of NOAA satellite data from December 1987 on the numerical analysis and forecast system of the People's Republic of China. This system has been described in Part 1, section 3, of this report.

The experiments run with the FGGE data set (described in part 1, Section 4), were reported with the December 1987 data, namely, NOSAT, SAT, ALLSAT and SATGUESS. An additional experiment was also performed. Named CONGUESS, it involves using the NOSAT analysis as the first guess, and then performing the ALLSAT experiment. We have decided, at this point, to only report on the NOSAT and SAT results from the December 1987 data sets. A report including the results of the ALLSAT, SATGUESS and CONGUESS experiments may be written in the near future with the help of our Chinese colleagues.

2. Data

The data used for this second experiment comes from NMC's operational data files (NMC Office Note 29). This set of data is more recent and should be better in quality than the FGGE data because the sounding retrieval procedures had undergone extensive improvements since 1979. The data period extends from December 14 through 28, 1987, and includes conventional as well as satellite data at both 00 and 12 GMT. The bulk of the conventional data are similar to that of the FGGE Global II-b level data used in the first set of experiments. The satellite data are unique in that soundings are available from both NOAA-9 and NOAA-10, but only

clear sounding data are available from NOAA-9.

Over the oceans, satellite data within a six hour window will generally give a good global coverage. For good global coverage over both land and ocean, in the case of the ALLSAT experiment, it was necessary to open the window to 9 hours. For each synoptic period in the data set, satellite soundings are available for up to plus six hours and minus eight hours. Thus all reasonable data windows could be accommodated by the data set used for all experiments. Figure 1 is the sample data coverage map for 12Z, December 16, 1987. As is evident, the data coverage is quite adequate. Note that no distinction is made between NOAA-9 and NOAA-10 data points.

Because there were two satellites available, there is a large overlap of data over most of the hemisphere. This raised concerns that the overlap may have caused averaging or blurring of some synoptic features in the analyses. This was tested by running experiments with and without overlapped data. We discovered that overlapped data from two satellites produced superior results over polar regions.

3. Forecast Impact

The impact of satellite soundings on Beijing Meteorological Center (BMC) model forecasts is measured objectively both in terms of error statistics of individual forecasts and in terms of the average or systematic errors taken over the entire test data set. The accuracy of NOSAT and SAT forecasts is determined by comparing them with appropriate NOSAT analyses in selected regions of good

radiosonde coverage, and computing rms errors and S1 skill scores. The selected verification areas are North Atlantic, East Pacific, West Pacific, Europe, West China, East China and North America. Here two parameters at Sea Level Pressure (SLP) and 500 mb height were selected for evaluation for 24 and 48 hour forecast periods. Beneficial impact is indicated by a reduction of the rms error and/or the S1 skill score for the SAT experiment compared with the NOSAT experiment.

3.1 S1 Skill Scores

Fig. 2 shows the temporal average of the 24 and 48 hour forecast S1 skill scores for SLP. The skill scores for 48 h forecast are generally lower than 80, which is the standard useful forecast skill score for SLP established by Halem et al., (1982). Over the oceans, a moderate positive impact of up to 6% was observed, but a disappointing negative impact of less than one percent, on the average, was observed over land. A similar result was reported by Thomasell et al., (1986), in which a 5.8% positive impact over ocean and a 0.7% negative impact over land was observed in their study using TIROS N and NOAA-6 data experimented with the numerical models of the Israel Meteorological Service. This current result appears to be consistent, given the fact that satellite data were used only over the oceans in the SAT experiment and no satellite data were used in the control experiment, NOSAT.

The 24 h sea level pressure S1 skill scores ranges from the 50's to the 60's with higher values over land than over oceans. Satellite impacts are mixed, and the results insignificant. These

skill scores are certainly larger than the comparable skillful level of 42 and 37 respectively for NOSAT and SAT reported by Kelly et al. (1978).

North America exhibits the poorest skill of all the verification areas and the largest negative impact for forecast periods of 48 h as seen from Fig 2. A time series of daily skill score for North America shows that SAT experiment runs are for the most part worse than those of NOSAT experiment (Fig. 3). The time series of skill scores for the Northern Hemisphere is shown in Fig. 4. The NOSAT skill scores are high, as was the case for North America. It is interesting, however, that the SAT experiment exhibits consistently a positive impact, in sharp contrast to the negative impact found in North America. Still, caution must be taken in the interpretation, because a well defined hemispherical truth is lacking.

The average S1 skill score for 48 h forecast of 500 mb height (Fig. 5) displays a pattern similar to Fig. 2 discussed earlier, but with only a meager 3% positive impact with the addition of satellite data over the oceans. Over land, however, the impact was mostly negative. The average forecast skill of 45, however, is significantly more skillful than 65, reported elsewhere.

West pacific stands out as the lowest skill score region in Fig. 5. A time series of skill score at each verifying data was plotted and compared with the skill score of the entire Northern Hemisphere shown in Fig. 6. It is indeed obvious that the skill scores at West pacific are substantially lower than the skill

scores of Northern Hemisphere which is around 40 for both NOSAT and SAT. This low skill score at West Pacific yields a 6.7% impact, substantially larger than the 5.5% averaged over ocean.

The 24 h forecast of 500 mb height in Fig. 5 shows in general negligible impact. The only exception is over West Pacific where the impact is 8%. The average skill scores over entire verification areas are 26.1 for both NOSAT and SAT.

3.2 Root-Mean-Square Differences

The root-mean-square error appears to be much more sensitive in response to the use of satellite data when compared with the S1 skill scores. At 24 h sea level pressure, some modest impact of satellite can be seen from several verification areas over both land and ocean. The large negative impact over West Pacific, however, is difficult to understand (Fig. 7), especially since skill score shows a slightly positive impact in Fig. 2. This is perhaps due to basic differences in nature between the skill score and rms errors are calculated. One may therefore not expect to see the same sign of the impact of satellite data on a forecast. (Chang, 1983). At 48 h the error increases but the overall pattern (Fig. 7) remains similar to the 24 h pattern seen also in Fig. 7. On the average, a modest 11.5% reduction of root-mean-square errors by satellite data was observed in the SAT experiment over land which compared with a 4.5% reduction over the oceans. The smaller reduction over the oceans, however, was influenced by a large error in the West pacific.

The 48 h root-mean-square errors of forecast at 500 mb height

for SAT experiment shown in Fig. 8 are generally below the useful threshold number of 64.8 listed in Table 1, Part I. An exception is over North America where the rms errors of SAT are around 80 and are not reduced by the addition of satellite data. The most encouraging result occurs, however, over West pacific where the rms errors of NOSAT are relatively smaller and reduction of errors by the addition of satellite data is 21% (Fig. 8). Large rms errors of NOSAT are also observed over East Pacific and Europe, but are reduced substantially by 13% and 22% respectively over these regions in response to satellite data. Throughout this experiment, North America stands out as one area where the experimental results are the least satisfactory. Figure 9 may provide some explanation, since SAT has, over most of the forecast period, larger rms errors. Moreover, the rms errors also appear to be more variable there, as evidenced by a larger standard deviation in Fig. 10. A very active synoptic pattern was observed over the United States continent between December 15 and 19, 1987. During this period, a nice wave pattern developed at 500 mb connecting two active low centers situated respectively over Northwest and Northeastern United States were maintained. A poor forecast during this period may contribute to the large rms errors over North America. If the root-mean-square errors over North America is discounted, the amount of reduction in rms errors should be around 8% for both land and ocean. This number is better than the impact of 2% reported by Desmarais et al. (1978) using VTPR and NIMBUS-6

data, and 4.3% by Druyan et al. (1987), also using VTPR sounding data.

Figure 8 also shows the 24 h root-mean-square error of forecast 500 mb height. Aside from displaying a smaller rms error, it also displays a pattern similar to those observed in the 48 h experiment. A point worth mentioning, however, is that there is roughly a 25 m difference in rms error between 24 and 48 h forecast based on the average Northern Hemispherical results. This brings the rms errors of 24 h forecast 500 mb height to around 42 and 39 m respectively for NOSAT and SAT, compared to 65 and 64 m for 48 h (Halem et al., 1978).

As we mentioned earlier, there is a great amount of data overlap over much of the area because of the inclusion of data from two satellites. Some concerns were raised as to the possibilities of blurring the synoptic features in the analysis due to excessive amount of data. This problem was examined by two parallel experiments. One series of experiments used all available satellite data. This results in multiple observations from overlapping orbits being used at individual gridpoints. In the other series an attempt was made to eliminate the overlapped data. In general, there was little difference in the two series except for the polar regions where the overlapped data produced superior results. An example is shown in Fig. 11 and 12 in which the root-mean-square errors appear to be almost identical.

The experimental results for the NOSAT and SAT are summarized in Table 1 for average forecast errors as well as S1 skill score

for 48 h forecast of sea level pressure and 500 mb height. The results are averaged over land and ocean verification areas. In addition, results are also computed respectively for unweighted averaging and weighted averaging by target size. It appears that the statistics changed a little depending on the size of the verification target, but was not significant enough to alter the direction of impact (not shown).

The verification statistics in Table 1 shows that, in general, that S1 skill score are not nearly as sensitive to satellite data as are the rms forecast errors since S1 skill score contains automatic penalties for attempts to hedge and hence a tendency to be overly conservative (Teweles and Wobus, 1954). In the rms forecast errors, positive impact by satellite data is consistent and significant both over land and oceans. The percent change of rms from NOSAT, respectively for land and ocean are 11.5% and 4.5% at sea level pressure, and 6.7% and 7.8% for 500 mb height. The S1 skill score on the other hand responds, favorably to the addition of satellite data by decreasing up to a maximum of only 5.5% over ocean. Over land, the responses are even smaller and negative, and are not compatible in sign with rms errors.

The 24 h forecast of rms errors and S1 skill score are smaller than their 48 h counterparts. The same is also true for the impact in general. It is possible that the smaller impact at 24 h is related to a shorter forecast time in which the opportunity for impact may not be sufficiently present.

3.3 Systematic Errors

The impact of satellite data on the systematic error of the BMC forecast model was also examined in the experiments using the December 1987 dataset. The systematic error is defined in the same way as in part 1, section 4. At this point only the systematic errors of the NOSAT and SAT experiments have been examined and, as in part 1, the statistical significance of differences in their respective systematic errors was determined using Student's t test.

Impacts upon model systematic error that will be discussed result from the SAT experiment runs, which allowed overlapping data from both NOAA satellites. The SAT runs using non-overlapping data resulted in some differing impacts in polar regions, but these were usually dominated by greater amounts of increased systematic error and therefore are not discussed.

For the reader's convenience, before detailing our current findings regarding the model's systematic error, we will list the major findings below.

1) Again, the BMC model is found to usually underforecast the strength of features in the mean sea level pressure (SLP) field, i.e., pressures are too low near high pressure systems and too high near low pressure systems.

2) Usually the same type of pattern is seen in the mean 500 mb height forecasts, i.e., the heights in troughs are not low enough and those in ridges are not high enough.

3) The use of satellite data (SAT experiment) in the BMC analysis/forecast system is found to have an impact upon the model's forecasts of sea level pressure and 500 mb height.

Systematic errors in the vicinity of surface lows over the oceans (especially over the N. Pacific) were found to be reduced somewhat. Errors associated with surface highs over land were also reduced, while the error for the one surface high observed over the ocean increased somewhat. A mix of positive and negative impacts was found in the 500 mb forecasts and will be detailed in a later section of this report.

3.4 Systematic Error of Sea Level Forecasts

Systematic errors of the 24, 48 and 72 h sea level pressure forecasts produced by the BMC forecast model were computed, as in part 1, as the difference between the mean forecast and mean verifying (NOSAT) analysis (forecast - analysis). Fig. 13 shows the mean verifying (mean NOSAT) sea level pressure analysis for the period 16-28 December 1987 at 1200 GMT. The mean SAT sea level pressure analysis is essentially the same as the NOSAT and is not shown. Fig. 14a-b show the mean 48 h forecast fields for this same period for the NOSAT and SAT experiments, while Figs. 15a-b show the corresponding systematic error fields.

In analyzing the model's systematic forecast errors from the December 1987 dataset we will focus on five features of the mean NOSAT sea level pressure analysis shown in Fig. 13. These are as follows:

1) Siberian High. The central pressure is 1036 mb and it is centered near 40N, 90E, with strong ridging (up to 1024 mb) extending over the North Pole and also eastward into the Pacific south of Japan.

2) Aleutian Low. The central pressure is 988 mb, with the center near 55N, 180E.

3) Eastern Pacific High. The central pressure is 1027 mb, located near 40N, 140W.

4) Icelandic Low. This is really a double-centered low with the deepest center (984 mb) located near 60N, 30W and the other center (995 mb) near 70N and 180E.

5) Western Mediterranean High. The central pressure is 1030 mb, centered near 30N and 5W. Pronounced ridging extends from the center westward across the Atlantic.

The computed systematic errors of the NOSAT and SAT 48 h forecasts of sea level pressure are shown in Figs. 15a-b. Several features in the error fields are similar to those described in part 1. These and other features arising from verification using the mean NOSAT sea level pressure analysis shown in Fig. 13 are described below by geographic region.

a. Eurasia

A large region of negative error associated with the Siberian High and Western Mediterranean High is seen in Figs. 15a-b, with maximum values of -16 and -14 mb respectively. Figs. 14a-b show that the NOSAT and SAT experiments each forecast the center of the Siberian high about 5 degrees south of the verifying analysis position and are too low with the central pressure by 4 and 3 mb respectively. This contrasts with the forecasts made using FGGE data which are discussed in Part 1 of this report. There, the mean position of the high's center was accurately forecast but the

central pressures forecast by all experiments were about 15 mb too low. The areas of large error near Lake Baykal produced by each forecast are due mainly to excessive troughing extending from the low northeast of Scandinavia. This strong troughing virtually cuts off the high near the north pole from the main high to the south.

Figs. 14a-b show the NOSAT and SAT experiments centering the Western Mediterranean High a bit too far south, with both mean 48 h forecasts being about 8 mb too weak. This weakens the pressure gradient between the high and the extension of the Icelandic Low to the northeast. In turn the breadth of this lobe of the Icelandic Low is greatly overforecast, resulting in the large areas of negative error centered over western Europe in Figs. 15a-b.

b. Pacific Ocean

Figs. 15a-b show areas of positive error centered southwest of the Aleutian Islands. The maximum error for both the NOSAT and SAT mean 48 h forecasts is +11 mb. In the case of the mean NOSAT forecast this error is caused by both the central pressure of the Aleutian Low being 7 mb too high and the mean position of the low being forecast too close to the Alaskan coast. Wallace and Woessner (1982), in describing the systematic error of the U.S. NMC 7LPE model, address that model's underforecasting of this feature's strength when describing a positive bias in 1000 mb height forecasts in the vicinity of the Aleutian low.

The mean 48 h SAT forecast of the low's central pressure is quite good (only 2 mb too high), but its similar positioning of the mean low too close to the Alaskan coast causes the positive error

to the southwest and also an area of negative error (maximum of -10 mb) centered over the Bering Strait.

The strength of the high over the easter Pacific seen in the mean NOSAT analysis (Fig. 13) is underforecast in both the NOSAT and SAT mean 48 h forecasts (Figs. 14a-b). Figs. 15a-b show the maximum errors to be -5 and -6 mb respectively, and in Fig. 15 we see that the mean SAT forecast does not show a closed high.

c. North Atlantic

While the position of the main low pressure system centered southwest of Iceland in the mean NOSAT analysis (Fig. 13) is forecast well in both the NOSAT and SAT mean 48 h forecasts, its strength is not. Both experiments forecast a central pressure 8 mb too high which helps lead to an area of positive error centered farther to the southwest of Iceland. NOSAT and SAT mean 48 h forecasts errors peak at +16 and +13 mb respectively in this region. Similarly, Wallace and Woessner (1982) find a significant positive bias in the U.S. NMC 7LPE model's forecast of 1000 mb heights in the vicinity of the Icelandic low. Neither mean forecast shows the pronounced trough curving to the southwest of the low in the NOSAT analysis (Fig. 13). This is directly reflected in the systematic error fields (Figs. 15a-b).

The systematic errors described above indicate (over this data period at least) that the BMC forecast model usually produces positive errors (SLP too high) in the vicinity of surface low pressure systems. Negative errors (SLP too low) are usually found in the vicinity of surface highs.

3.5 Systematic Error of 500 mb Height Forecasts

Systematic errors of the 24, 48 and 72 h 500 mb height forecasts produced by the BMC forecast model have been computed in the same manner as described earlier for sea level pressure. Figs. 16a-b show the mean NOSAT and SAT 500 mb analyses for the period 16-28 December 1987 at 1200 GMT. As was the case with the FGGE dataset, significant differences between the analyses are difficult to spot when visually comparing them, so NOSAT 500 mb heights were subtracted from SAT heights and presented as the difference map shown in Fig. 17. The differences resulting from including satellite data over the oceans in the SAT experiment are clearly evident.

The mean 48 h 500 mb height forecasts for the two experiments are shown in Figs. 18a-b and the corresponding systematic error fields are shown in Figs. 19a-b. Descriptions of the main features in the systematic error fields are presented below by geographic region.

a. Eurasia

The systematic error maps for the NOSAT and SAT 500 mb height forecasts (Figs 19a-b) show similar regions of negative error centered southwest of Lake Baykal (maximum errors of -111 and -101 m respectively). This stems from neither forecast (Figs. 18a-b) showing the ridging seen in the NOSAT analysis (Fig. 16a). The ridge axis extends along 85E roughly between 40 and 60N. The two forecast fields actually seem to show a slight trough at this location.

Another region of negative error, produced by both forecasts, is centered near 50N and 10E (maximum errors are -105 and -90 m respectively). This appears to be due to neither mean 48 h forecast showing strong enough ridging over western Europe. The mean NOSAT analysis (Fig. 16a) shows greater heights in the ridge with the axis roughly along 5E. The mean forecast fields (Figs. 18a-b) show weaker ridging and instead align the axis more along the prime meridian. Wallace and Woessner (1982) note that the U.S. NMC 7LPE model has been found to show a similar negative bias in forecasting 500 mb heights in standing ridges over western Europe.

The SAT systematic error map (Fig. 19) shows an area of positive error near 50N and 130E that is barely indicated in the respective NOSAT map. The maximum error is +75 m and seems to be caused by the mean 48 h SAT forecast not producing the troughing seen through Manchuria in the NOSAT analysis (Fig. 16a).

Neither model forecast builds the cutoff high near 80 N and 140E enough, resulting in the areas of negative error seen there in Figs. 19a-b.

b. Pacific Ocean

Figs. 19a-b each show an area of negative error lying south of Japan. The NOSAT forecast produces a much larger error; 96 m compared to -39 m for SAT. This appears to be the product of the nearly zonal flow just south of Japan along 140E in the NOSAT analysis (Fig. 16a) being replaced by the troughing which is especially evident in the mean NOSAT forecast (Fig. 18a). The area of negative error extends well to the east for the mean NOSAT

along 45W. Both model runs do a good job of forecasting the mean positions of the trough axes, but neither run deepens the troughs enough. Similar positive biases in troughs over the east coasts of continents are generally produced by the U.S. NMC 7LPE model (Wallace and Woessner, 1982).

3.6 Differences in Systematic Errors of Sea Level Pressure Forecasts

As in part 1, to determine any differences between systematic errors from the experimental runs which used satellite data and the NOSAT experiment, the difference between the absolute values of the systematic errors from the 48 h mean SLP forecasts were analyzed. Fig. 20a shows isopleths of $E_{SAT} - E_{NOSAT}$. Negative values imply a reduction in systematic error, positive values an increase in systematic error. Heavy dashed lines show areas where E_{SAT} is significantly different from E_{NOSAT}.

Fig. 20 shows several regions where the use of satellite data over the oceans has a significant impact on the mean 48 h sea level pressure forecast. The two largest areas are found in Eurasia and each are associated with a reduction in systematic forecast error. The use of satellite data has apparently lessened the amount by which the Western Mediterranean and Siberian Highs were underforecast.

Areas of reduced systematic error also are coincident with the mean positions of the Icelandic and Aleutian Lows, however in each case the areas found to be statistically significant comprise only a small portion of the area affected.

Areas of increased systematic error are also seen in Fig. 20a. The area near the Bering Strait was not found to be statistically significant, but three other areas (eastern North America, the eastern and western Pacific) were. In the eastern Pacific, the use of satellite data has been slightly detrimental in forecasting the surface high shown in Fig. 13. Near Japan, the surface ridging seen in the SAT mean 48 h forecast (Fig. 14b) has combined with the inaccurate placement of the Aleutian low to cause a large area of significantly increased systematic error. Both the NOSAT and the SAT mean 48 h forecasts (Fig. 14a-b) had trouble with the extent of troughing reaching southwestward from the Icelandic Low. Both forecasts underplayed this trough and built too strong of a high to the south. The SAT run did this to a greater extent, resulting in the increased error seen off the east coast of North American (Fig. 20a).

Like the results from the FGGE data period described in part 1, the results from the December 1987 period show satellite data having significant impact on the systematic error of the model's sea level pressure forecasts. Using the data over the oceans was again found to have a mainly beneficial impact on the mean SLP forecast near the center of well-developed oceanic surface lows. However, in the wakes of these systems, to the southwest, systematic error significantly increased because the strength of troughing was underforecast.

The most impressive beneficial impacts were seen in the reduction of forecast error near surface highs over land. There

is some evidence from this data period however that satellite data use over the oceans was a detriment in forecasting the strength of oceanic surface highs.

3.7 Differences in the Systematic Error of 500 mb Height Forecasts

Differences in the systematic error of 500 mb height forecasts were calculated according to the method described in section 3.6 for sea level pressure forecasts. The resulting field is presented in Fig. 20b. As before, negative values imply a reduction in systematic error, positive values an increase in systematic error and heavy dashed lines denote areas where E_{SAT} is significantly different from E_{NOSAT}.

In Fig. 20b we see several regions where using satellite data over the oceans has had significant impact on the mean 48 h 500 mb height forecast. A large region of beneficial impact is seen over Europe where the mean 48 h SAT forecast (Fig. 18b) has done a better job in building the ridge than the mean 48 h NOSAT forecast (Fig. 18a). The same type of effect is observed near Lake Baykal in Fig. 21b due to the SAT runs doing a better job than the NOSAT runs with the ridge in that vicinity.

Continuing to the east in Fig. 20b, we see a pattern very similar to that in the $E_{SAT} - E_{NOSAT}$ map from the FGGE data period. Areas of increased error are seen near Manchuria and the Yellow Sea while a large area of decreased error lies mainly south of Japan. The increased error is caused by the SAT forecast (Fig. 18b) showing less of the troughing analyzed through Manchuria in Fig. 10a than the NOSAT forecast (Fig. 18a).

The mean SAT forecast (Fig. 18b) builds an especially pronounced bogus ridge along the west shore of the Yellow sea just as it did in the FGGE data runs. But again, the SAT forecast is less likely to dig a nonexistent trough in the flow south of Japan than the NOSAT forecast. This leads to the large area of decreased error seen there in Fig. 20b. The last area of statistically significant impact in Fig. 20b is found over eastern North America. This area of increased error resulted from the SAT runs being weaker in digging the trough there than the NOSAT runs.

While not proving statistically significant, the area of increased error seen in the north Pacific in Fig. 20b is interesting. In the FGGE data period the mean 500 mb analysis in that region was very similar to that of the December 1987 analysis, yet in that experiment an area of significantly reduced error was found in approximately the same location.

The results presented above indicate that using satellite data over the oceans impacts the systematic error of the model's 500 mb height forecasts. The impact on the forecasts along the east Asian coast for this data period was found to be remarkably similar to that observed in the FGGE data experiment. The mean analysis in each case showed a closed low over eastern Siberia with weak troughing extending southwestward just inland from the Pacific coast. Off the coast the flow was highly zonal. The SAT runs with each dataset tended to be inferior to the NOSAT runs at predicting the slight trough just inland, i.e., the troughing was too weak. Just off the coast however, the mean SAT forecast was superior to

the NOSAT as it was less likely to form a trough where zonal flow had been analyzed.

It is difficult, using the December 1987 data, to make general conclusions elsewhere regarding the effects of satellite data on the 500 mb forecasts. Satellite data was useful in forecasting the mean ridge over western Europe but was detrimental in forecasting the mean trough over eastern North America.

4. Discussion and Conclusion

The second impact experiment used the December 1987 data set which were considered to be of good quality because in that data set soundings were available from two satellites in which the sounding retrieval procedures had undergone extensive improvements since 1979. Nevertheless, it failed to yield results commensurate with the expectation of a better quality data set.

In the SAT experiment, where satellite soundings are combined with equal weighting with radiosonde data, a maximum positive impact on rms errors of up to 11.5% was observed over land at SLP and 8% over ocean at 500 mb height for forecast periods of 48 hours. This figure of 11.5% is comparable to Kelly's result (1977), but still less than the 14% reduction of rms errors effected by the satellite data over the ocean for SLP reported by Thomasell et al., (1986). However, the present results represent a substantial improvement over the impacts of 6.5% and 3% reported respectively by Halem et al. (1978) and Druyan et al. (1978) based on their 48 h results over land. The S1 skill score appears to respond less favorably to satellite data. A meager S1 score

increase of 5.5% was observed at SLP for 48 h forecast over ocean. This impact is considered to be weak and insignificant.

Overall, the verification statistics shows that there is more impact for forecast periods of 48 hours than for 24 hours. This is true for both S1 skill score and rms error, particularly over the oceans. Certainly, it is highly desirable to see the satellite data applied over ocean moving with the system can produce a positive impact over land eventually, given sufficient amount of forecast time. In this respect, Europe's favorable response to satellite data may well illustrate this point, despite a poor response over North America.

The use of the December 1987 satellite data over the oceans is found to have significant impact on the systematic error of the BMC model's sea level pressure and 500 mb height forecasts. Reductions in error were found near well-developed oceanic surface lows, but errors were increased in troughing to the southwest of the lows. No clear impact was observed upon the 500 mb troughs associated with these lows.

Satellite data were also found to reduce forecast error in the vicinity of broad surface high pressure systems found over Eurasia. Forecast error was also reduced near the associated 500 mb ridges.

A particularly interesting finding relates to the very similar upper air patterns found over the East Asian coast in the FGGE and December 1987 mean 500 mb analyses which were described in section 3.7. In each case the introduction of satellite data over the oceans correlates with an increase in systematic error along the

coast and a decrease in systematic error just off the coast.

The December 1987 data set has two satellite soundings available. One series of experiments used all available satellite data. This results in multiple observations from overlapping orbits being used at individual gridpoints. In the other series, an attempt was made to eliminate the overlapped data. In general, there was little difference found in the two series except for the polar regions where the overlapped data produced superior results.

Basic emphasis of December 1987 experiments was focused on the analysis of results from the SAT experiment and the control experiment, NOSAT, that has no satellite data. The other experiments, namely, ALLSAT, SATGUESS, and CONGUESS were found to yield unexpected results and require considerably more study before we can be confident of understanding and explaining them.

Table 1. Skill score and rms forecast errors averaged over the December 1987 dataset for the NOSAT and SAT experiments.

		Land			Ocean		
Parameter	Forecast period (h)	NOSAT	SAT	Percent change from NOSAT	NOSAT	SAT	Percent change from NOSAT
				S1 skill score			
SLP	24	59.9	59.3	-1	53.8	52.2	-3.0
(%)	48	76.5	76.8	0.4	76.6	72.4	-5.5
500 mb z	24	26.3	26.5	0.8	25.8	25.6	-0.8
(%)	48	37.6	38.5	2.4	40.5	39.3	-3.0
				RMS forecast error			
SLP	24	5.5	5.0	-9.1	5.3	5.3	0.0
(mb)	48	8.7	7.7	-11.5	8.8	8.4	-4.5
500 mb Z	24	39.6	37.1	-6.3	44.8	41.1	-8.3
(m)	48	67.3	62.8	-6.7	68.2	62.9	-7.8

References

Chang, J.T., 1983: Summary Report of preliminary analysis for impact of satellite sounding data on IMS Model Forecasts. Systems and Applied Sciences Corporation, contract #NA-81-SAC-00725.

Desmarais, A., S. Tracton, R. McPherson, and R. van Haaren, 1978: The NMC Report on the Data Systems Test. NASA Contract S-70252-AG, National Meteorological Center, Camp Springs, Md., 313 pp.

Druyan, L.M., T. Ben-Amran, Z. Alperson, and G. Ohring, 1978: The impact of VTPR data on numerical forecasts of the Israel Meteorological Service. Mon. Wea. Rev., 106, 859-869.

Halem, M., M. Ghil, R. Atlas, J. Susskind, and W.J. Quirk, 1978: The GISS sounding temperature impact test. NSAS Tech. Memo. 78063, NSAS Goddard Space Flight Center, Greenbelt, Md., 421 pp.

Halem, M., E. Kalnay, W.B. Baker and R. Atlas, 1982: An assessment of the FGGE satellite observing system during SOP-1. Bull. Amer. Meteor. Soc., 63, 407-426.

Kelly, G.A.M., 1977: A cycling experiment in the Southern Hemisphere using VTPR data. Australian Numerical Meteorology Research Center, Melbourne. (unpublished manuscript).

Kelly, G.A.M. and W.L. Smith, 1978: Impact of Nimbus-6 temperature soundings on Australian region forecasts. Bull. Am. Meteorological Soc., 59, 393-405.

Teweles, S. Jr., and H.B. Wobus, 1954: Verification of prognostic charts. Bull Amer. Meteor. Soc. 35, 455-465.

Thomasell, A. Jr., A. Gruber, H. Brodrick, N. Wolfson and Z. Alperson, 1986: The impact of satellite soundings on the Numerical Forecasts of the Israel Meteorological Services. Mon. Wea. Rev. 114, 1251-1262.

Wallace, J.M. and J.K. Woessner, 1982: Analysis of forecast errors in the NMC hemispheric primitive equation model. Mon. Wea. Rev., 109, 2444-2449.

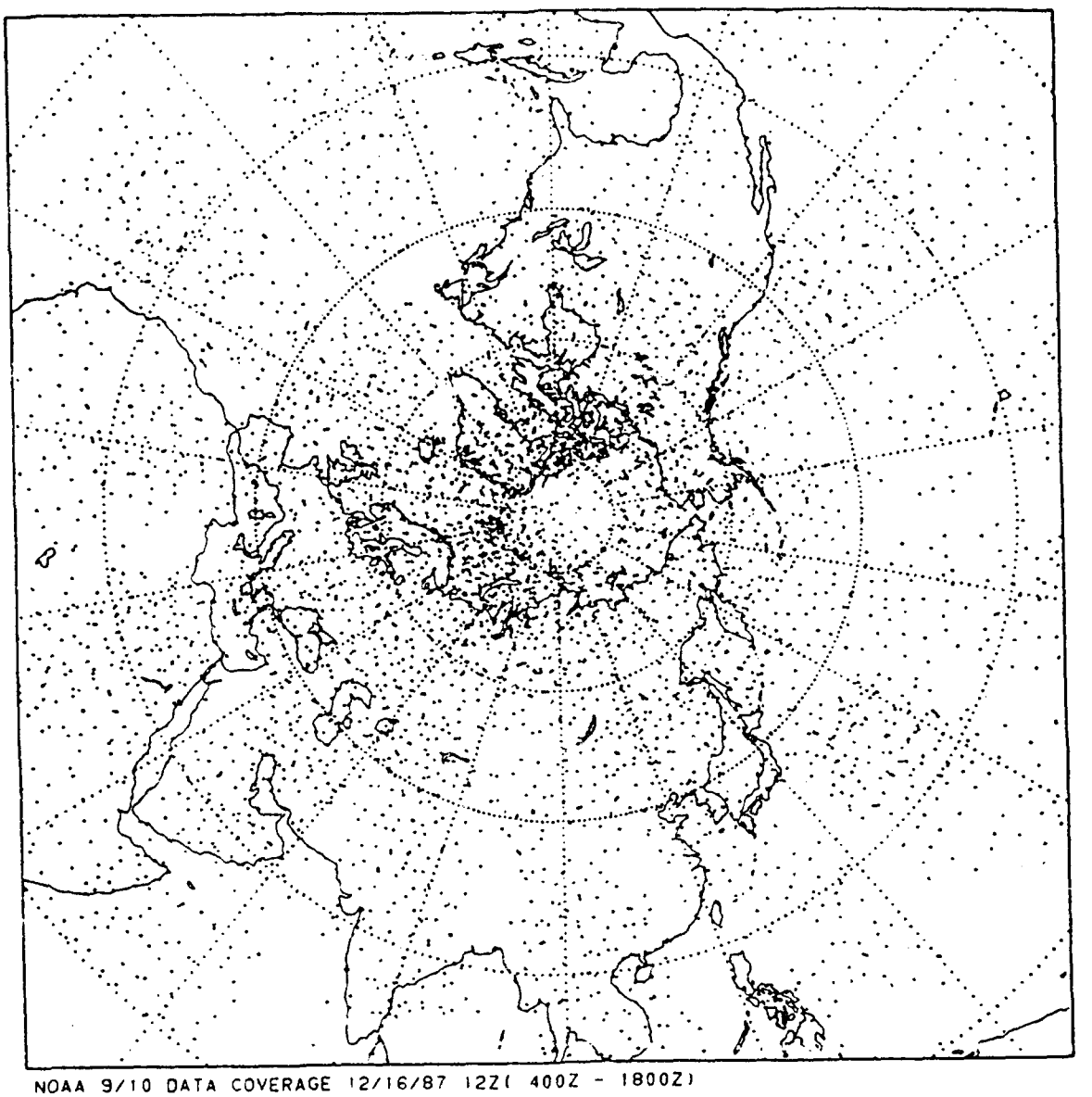

Fig. 1. Sample Data Coverage of both NOAA 9 and 10 at 12 Z, Dec. 16, 1987.

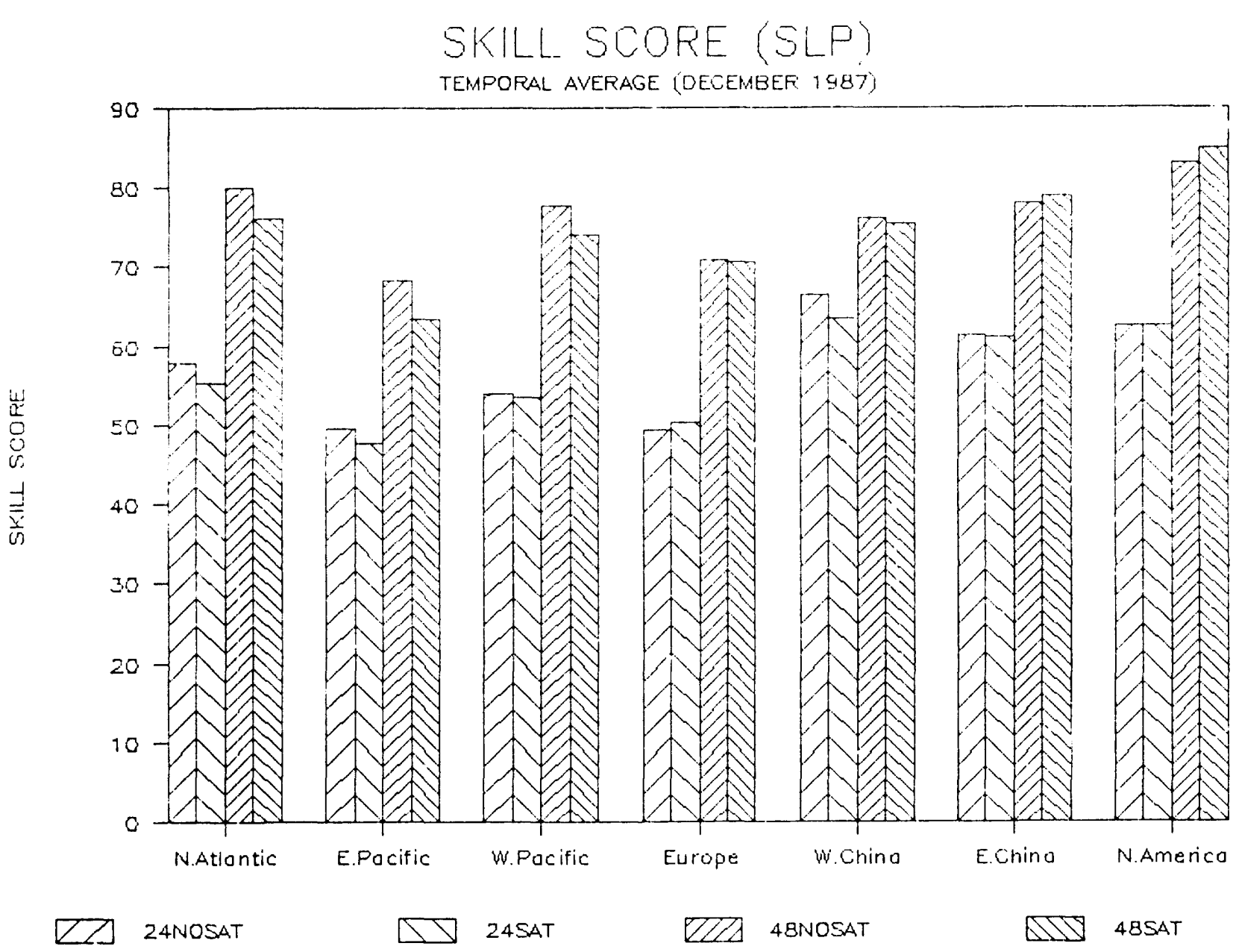

Fig. 2. The average skill score of 24 and 48-hour forecasts of Sea Level Pressure (SLP) with and without satellite data (SAT and NOSAT) over each verification area.

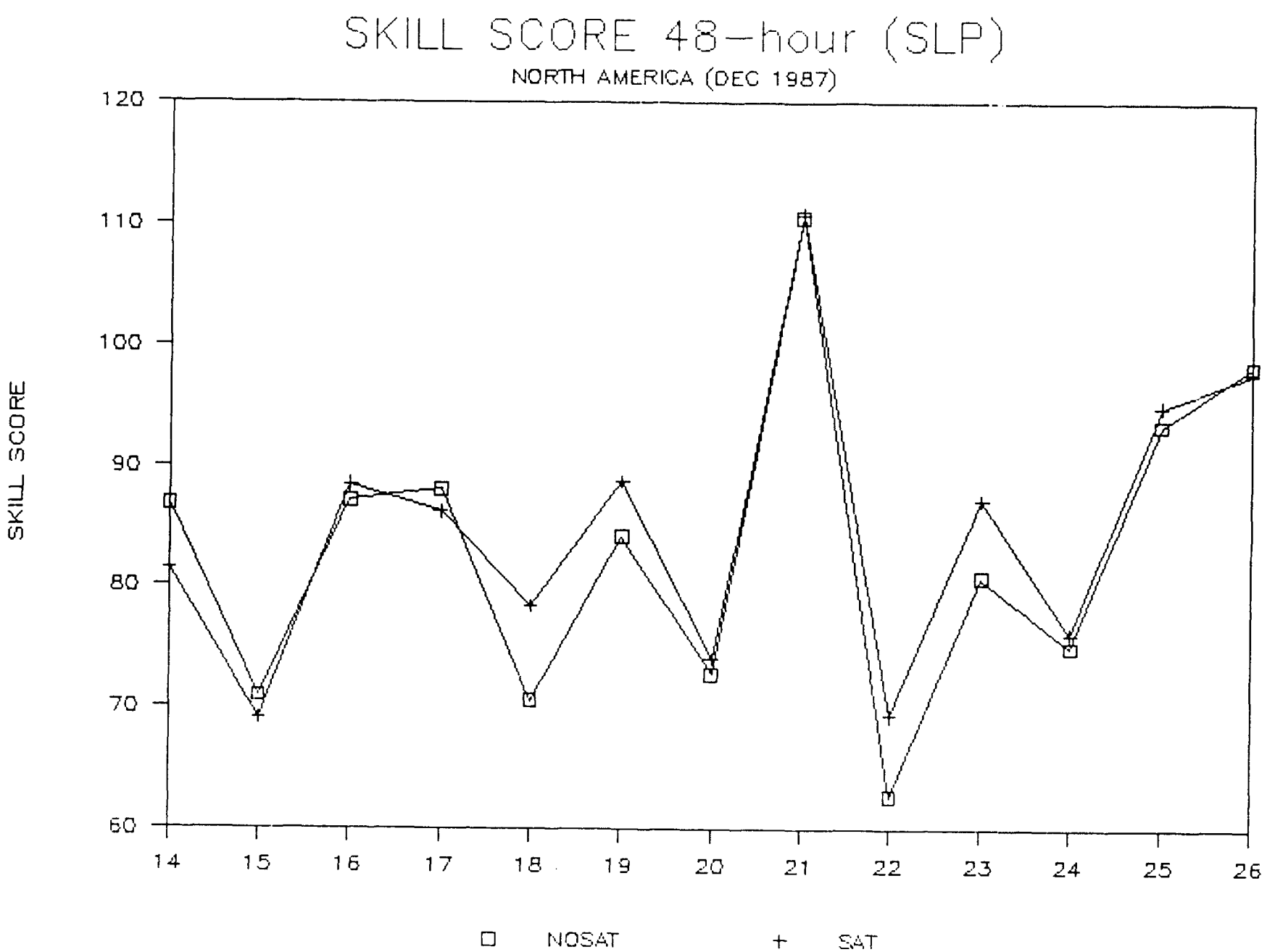

Fig. 3. Daily skill score of 48-hour forecast of SLP for NOSAT and SAT over each verification area.

Fig. 4. Same as in Figure 3 but for Northern Hemisphere.

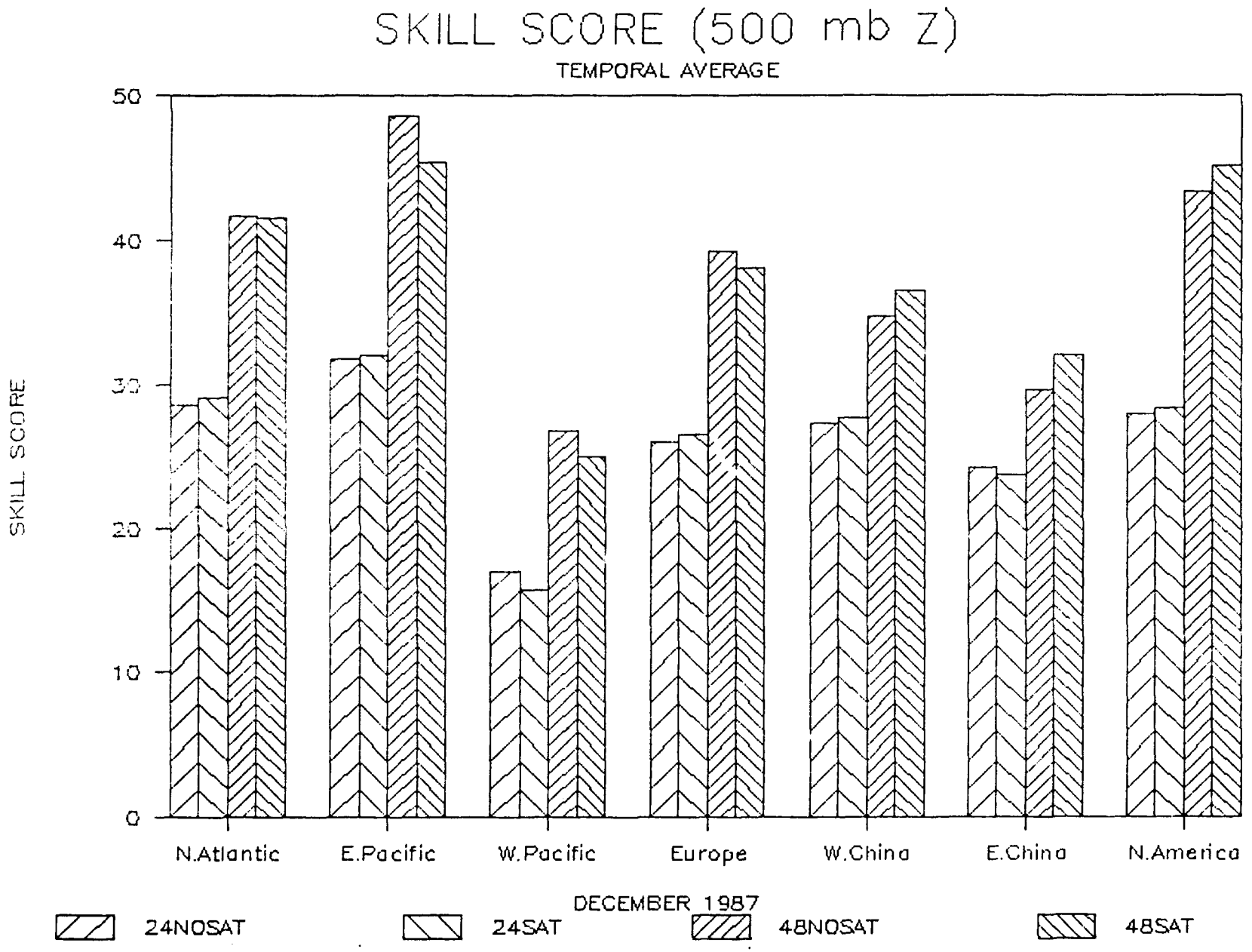

Fig. 5. The average skill score of 24 and 48-hour forecasts of 500 mb height over each verification area for NOSAT and SAT experiments.

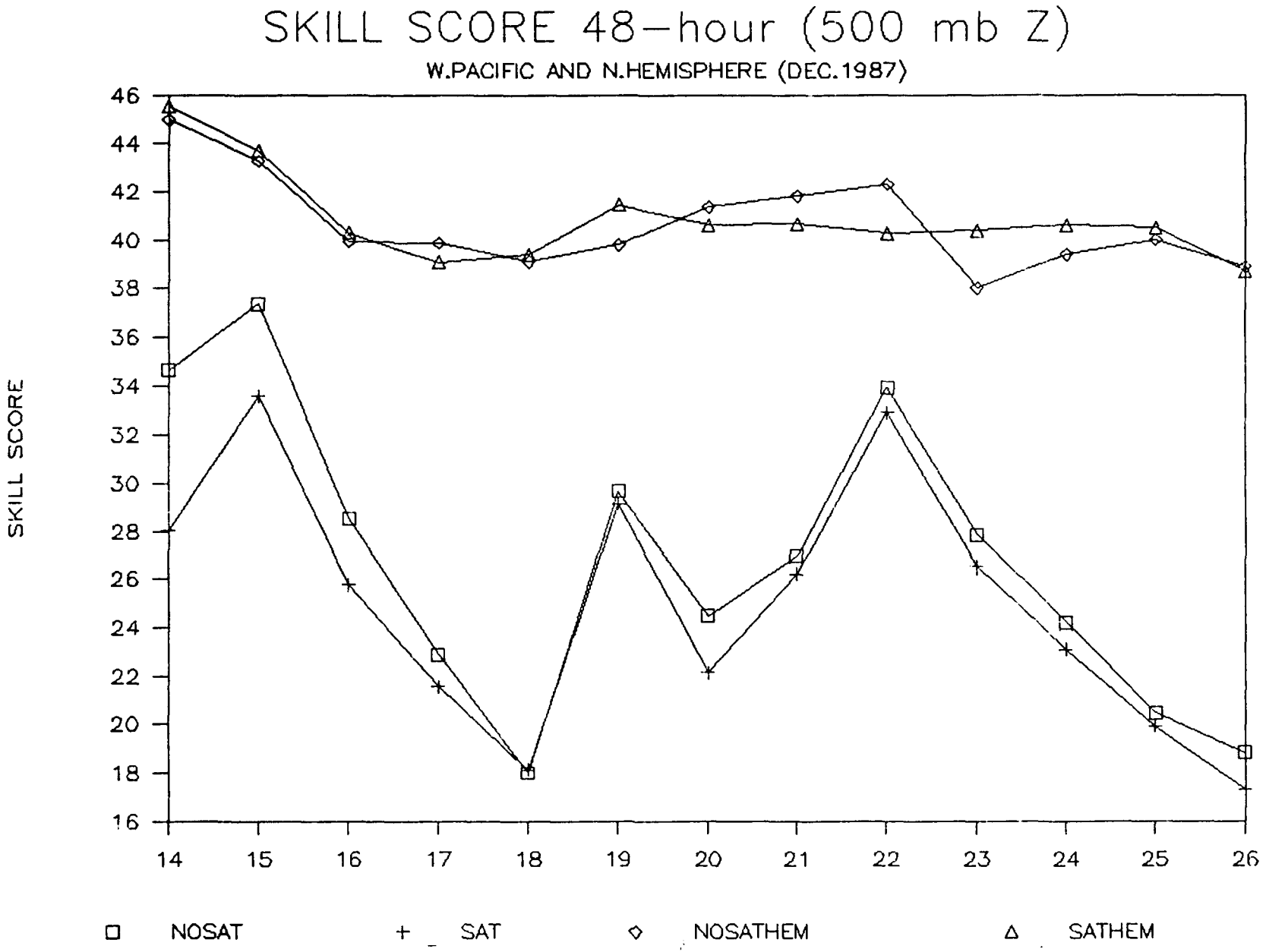

Fig. 6. Daily skill score of 48-hour forecast of 500 mb height over West pacific and Northern Hemisphere, where NOSATHEM and SATHEM mean respectively the NOSAT and SAT experiments over Northern Hemisphere.

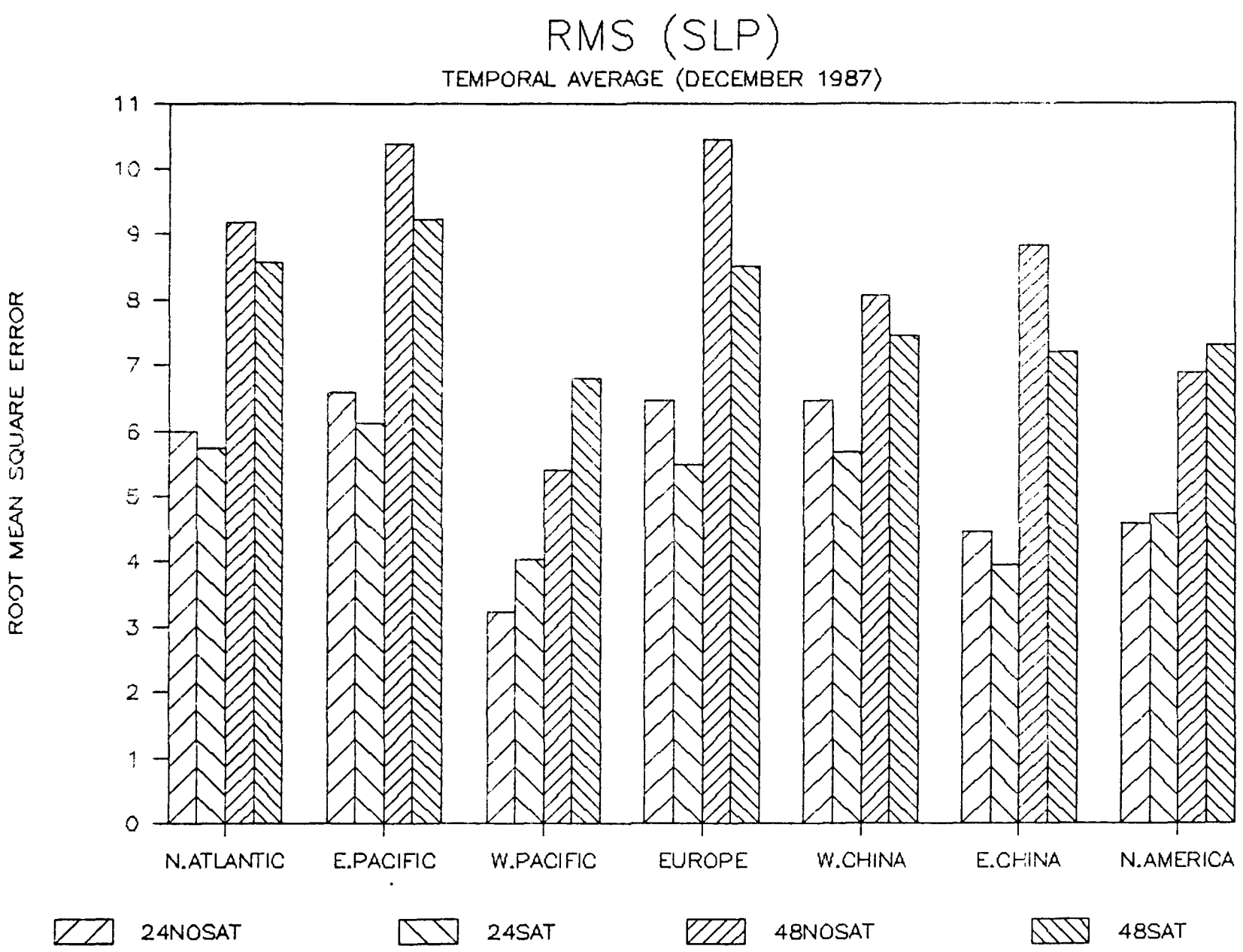

Fig. 7. **Average root mean square error 24 and 48-hour forecasts of seal level pressure for NOSAT and SAT experiments over each verification area.**

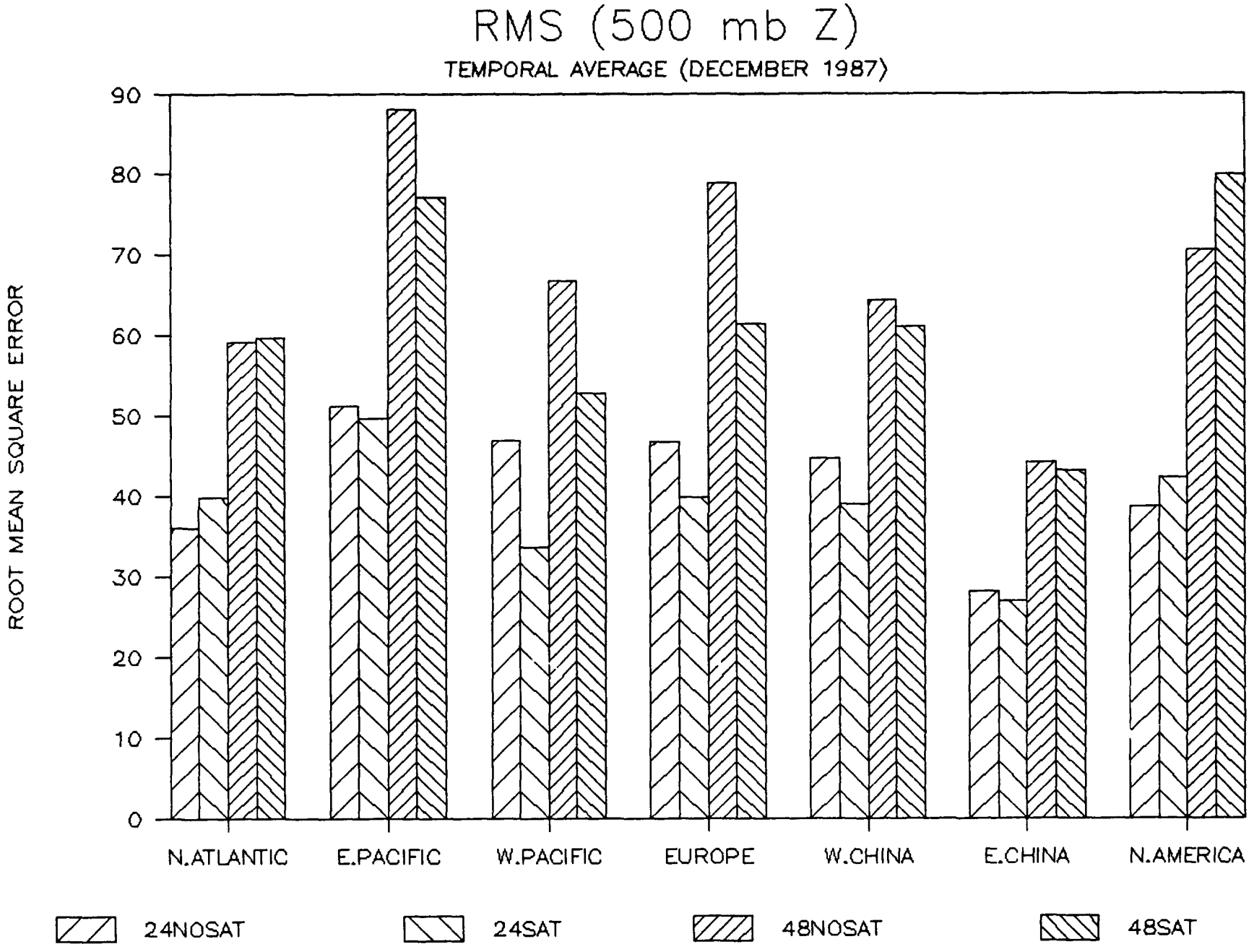

Fig. 8. Same as in Fig. 7 but for 500 mb height.

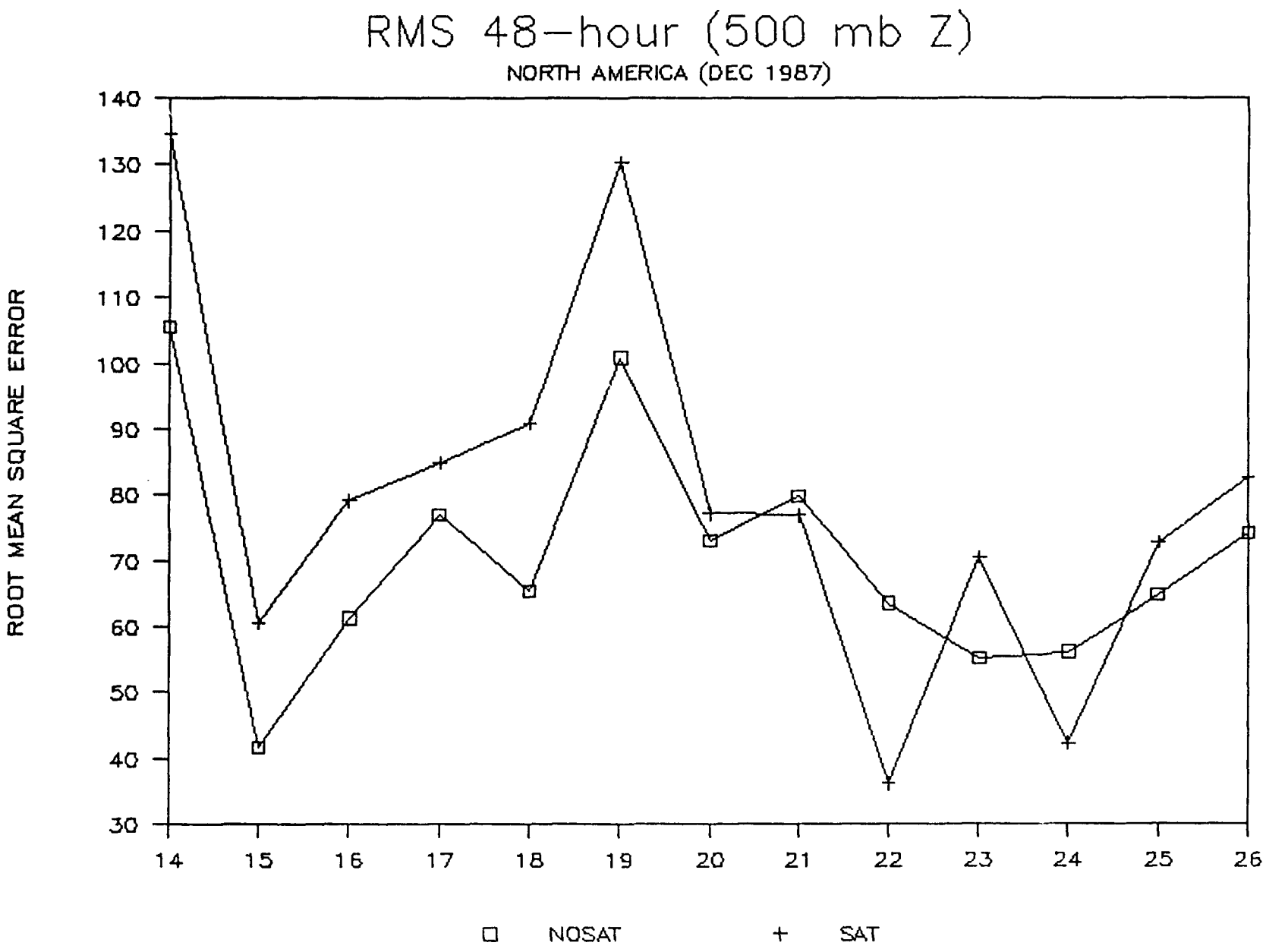

Fig. 9. Daily root mean square error of 48-hour forecast of 500 mb height over North America for NOSAT and SAT experiments.

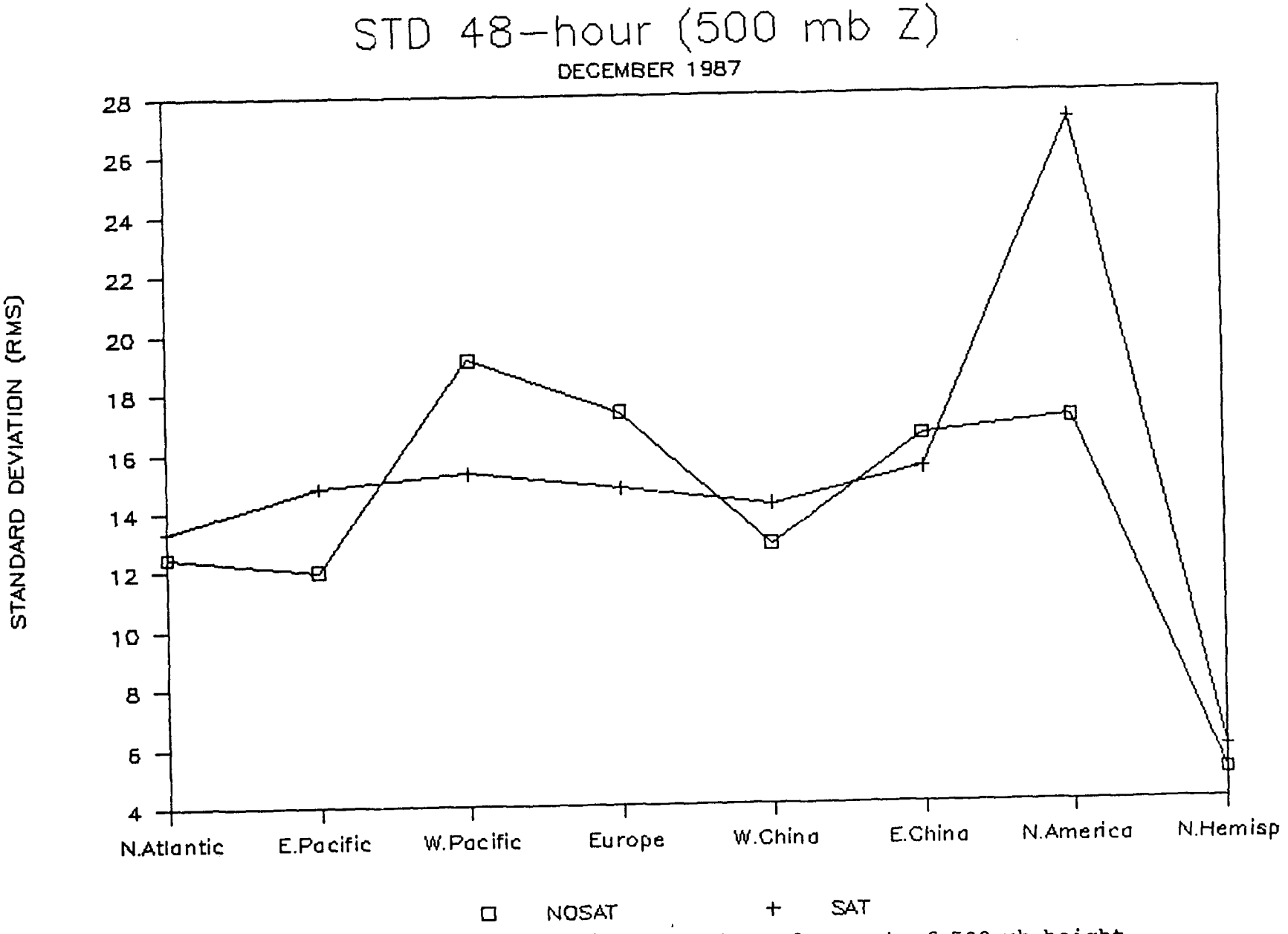

Fig. 10. Standard derivation of 48-hour forecast of 500 mb height over each verification area for NOSAT and SAT experiments.

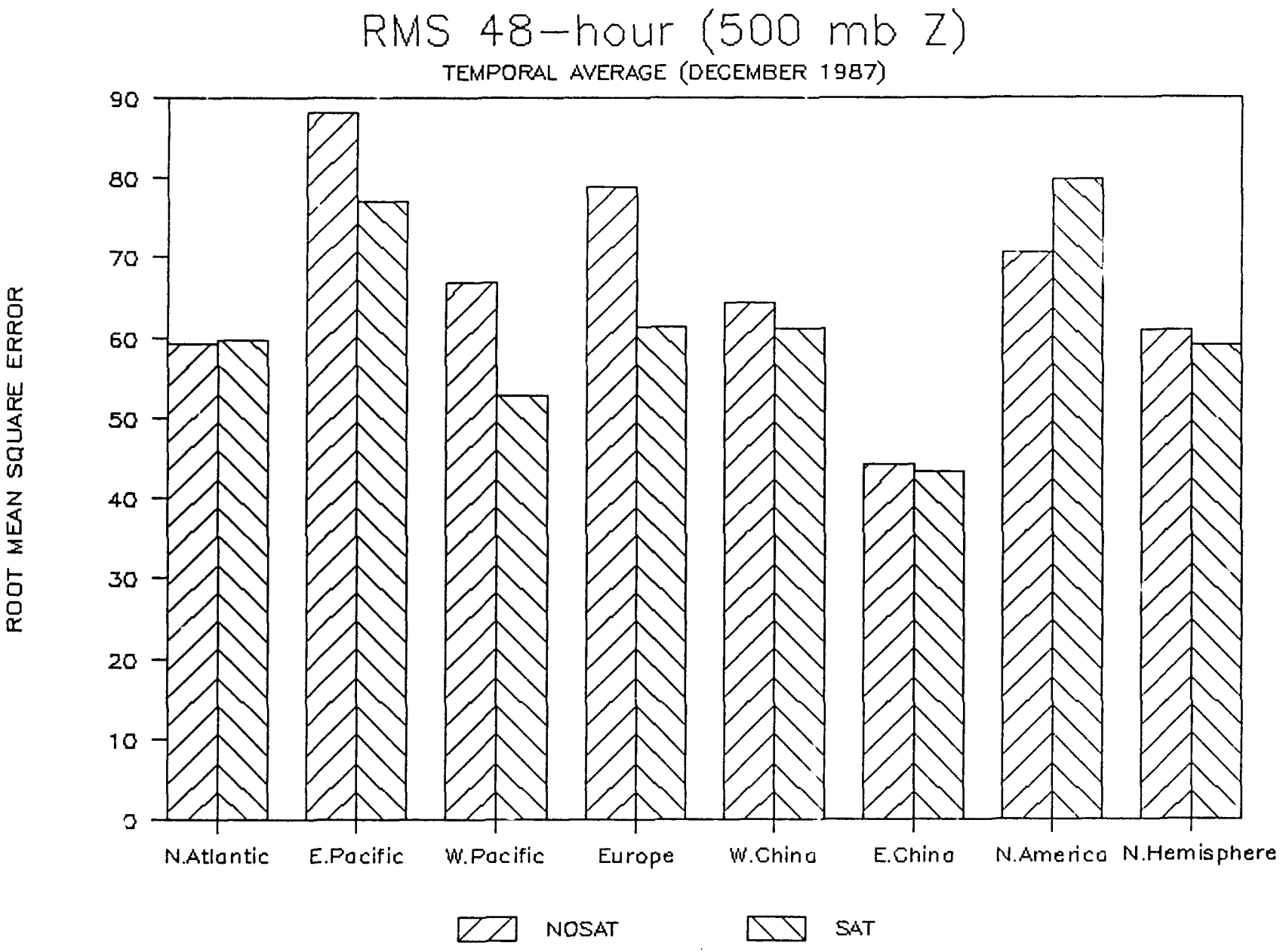

Fig. 11. Average root mean square error of 48-hour forecast of 500 mb height for NOSAT and SAT experiments over each verification areas, including the entire Northern Hemisphere.

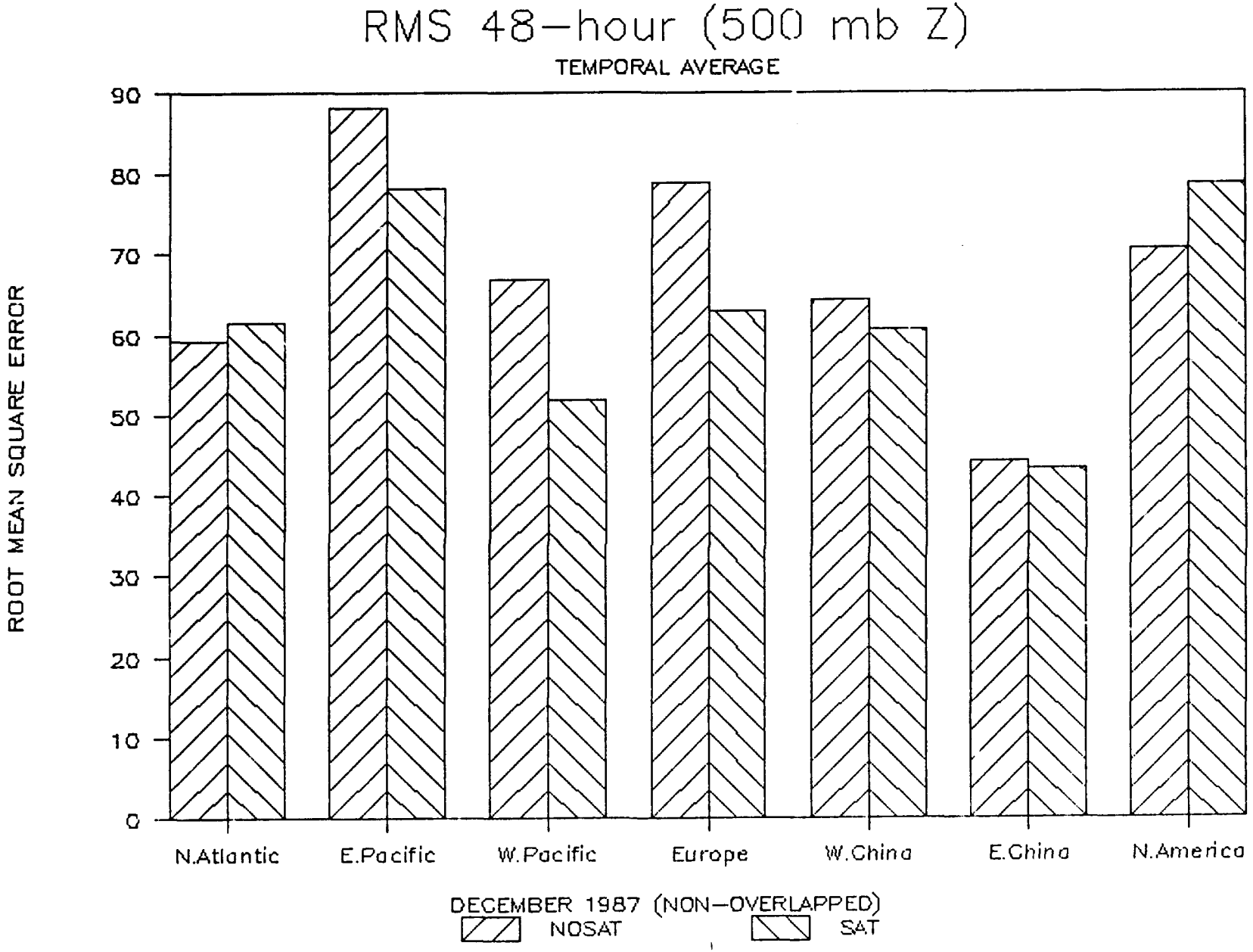

Fig. 12. Same as in Figure 11 but for using only those soundings closed to 1200 GMT in regions where soundings at different times exist.

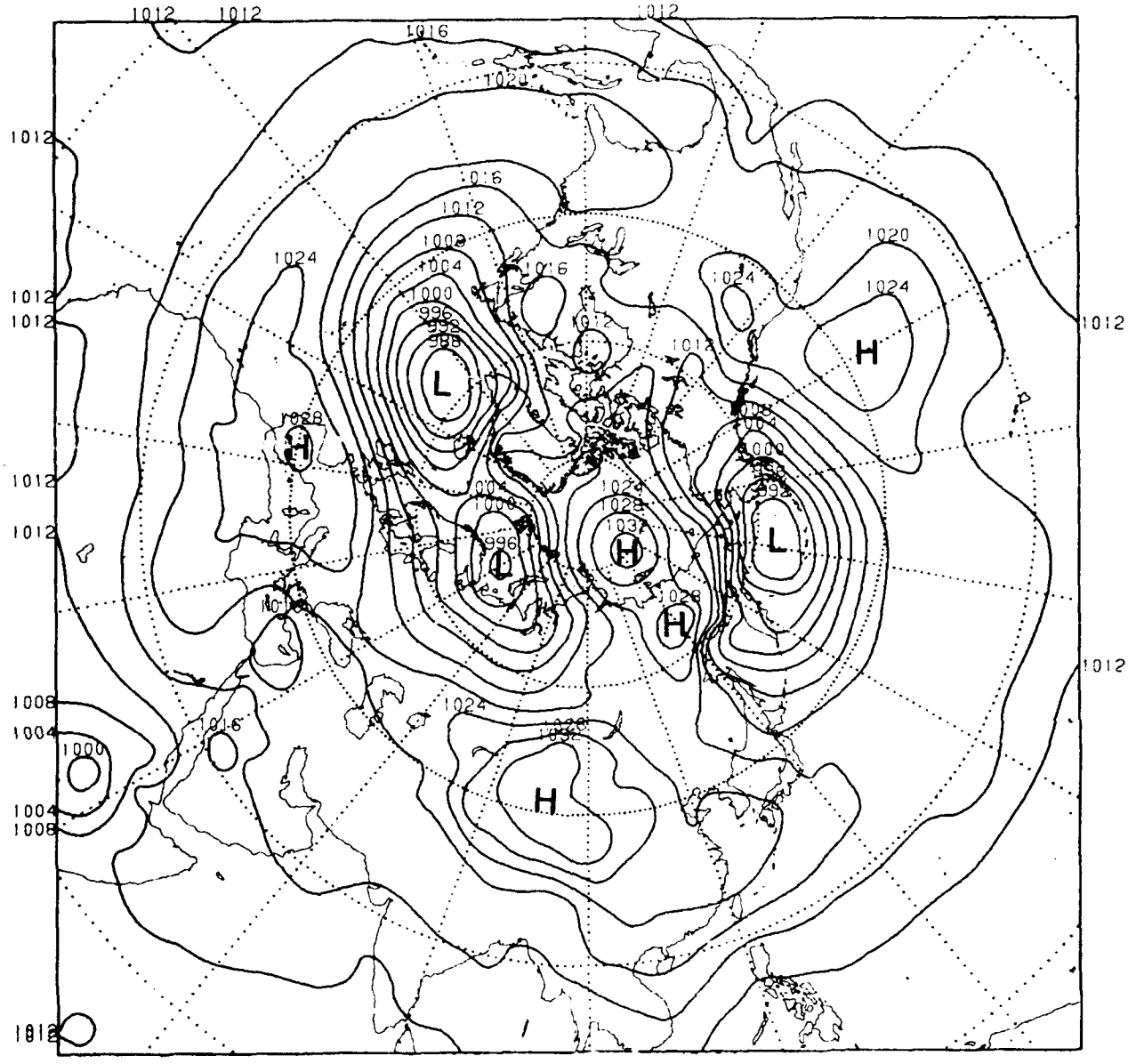

Fig. 13. Average NOSAT sea level pressure analysis for 16-28 December 1987 at 1200 GMT.

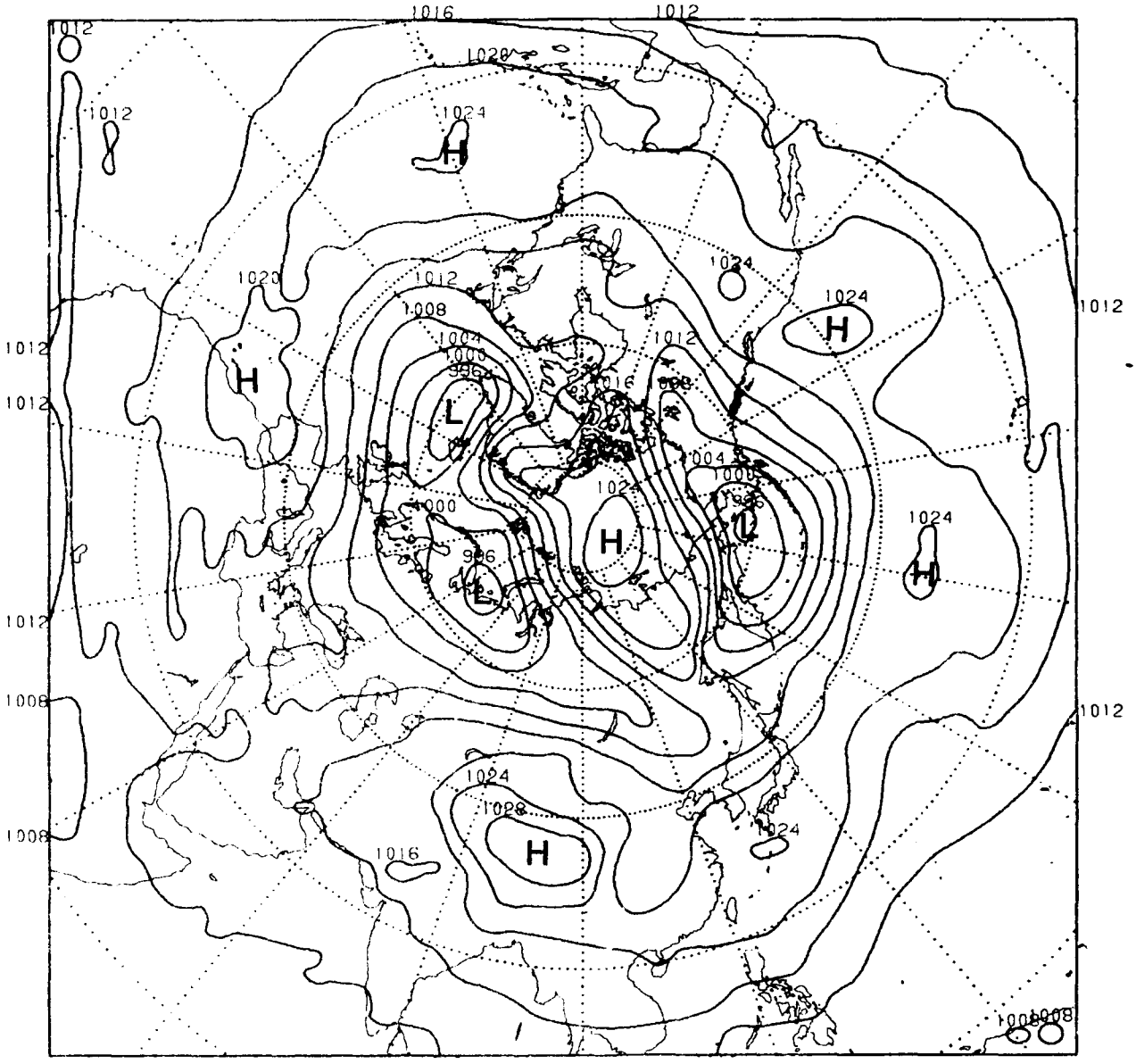

Fig. 14a. Average 48 hour NOSAT forecast of sea level pressure for 16-28 December 1987 at 1200 GMT.

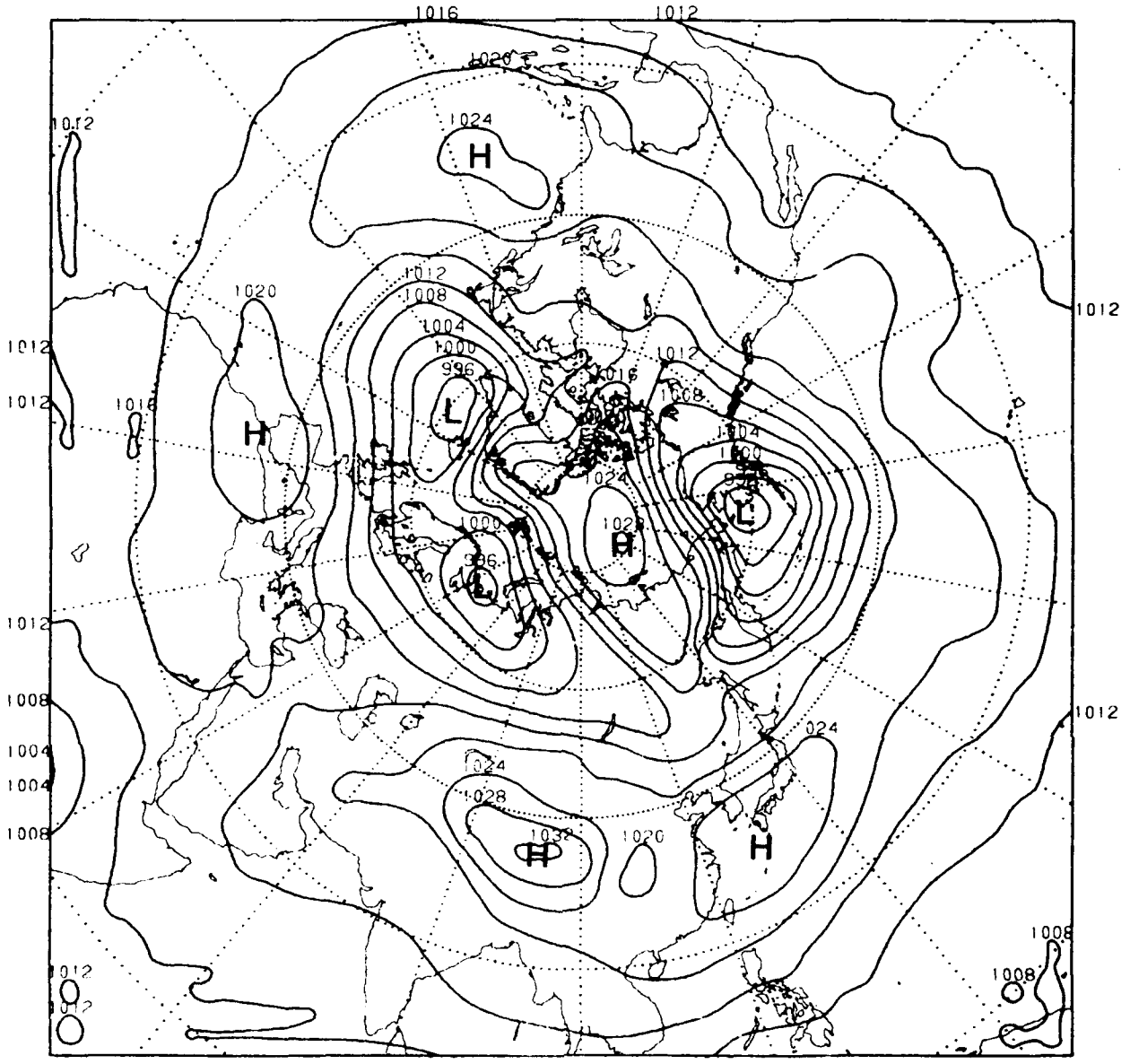

Fig. 14b. Average 48 hour SAT forecast of sea level pressure for 16-28 December 1987 at 1200 GMT.

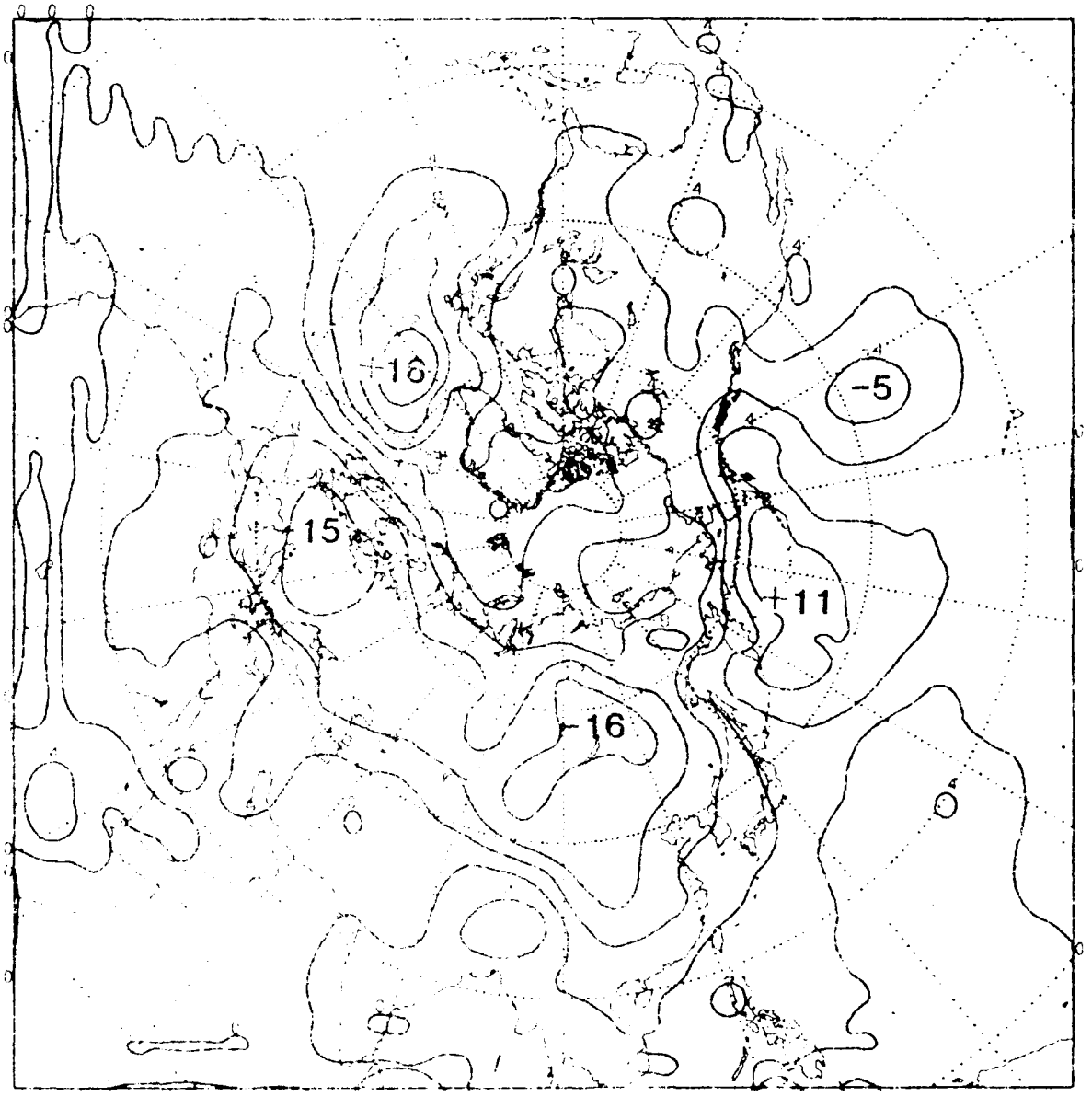

Fig. 15a. Systematic error (mb) of the NOSAT 48 hour forecast of sea level pressure for 16-28 December 1987 at 1200 GMT.

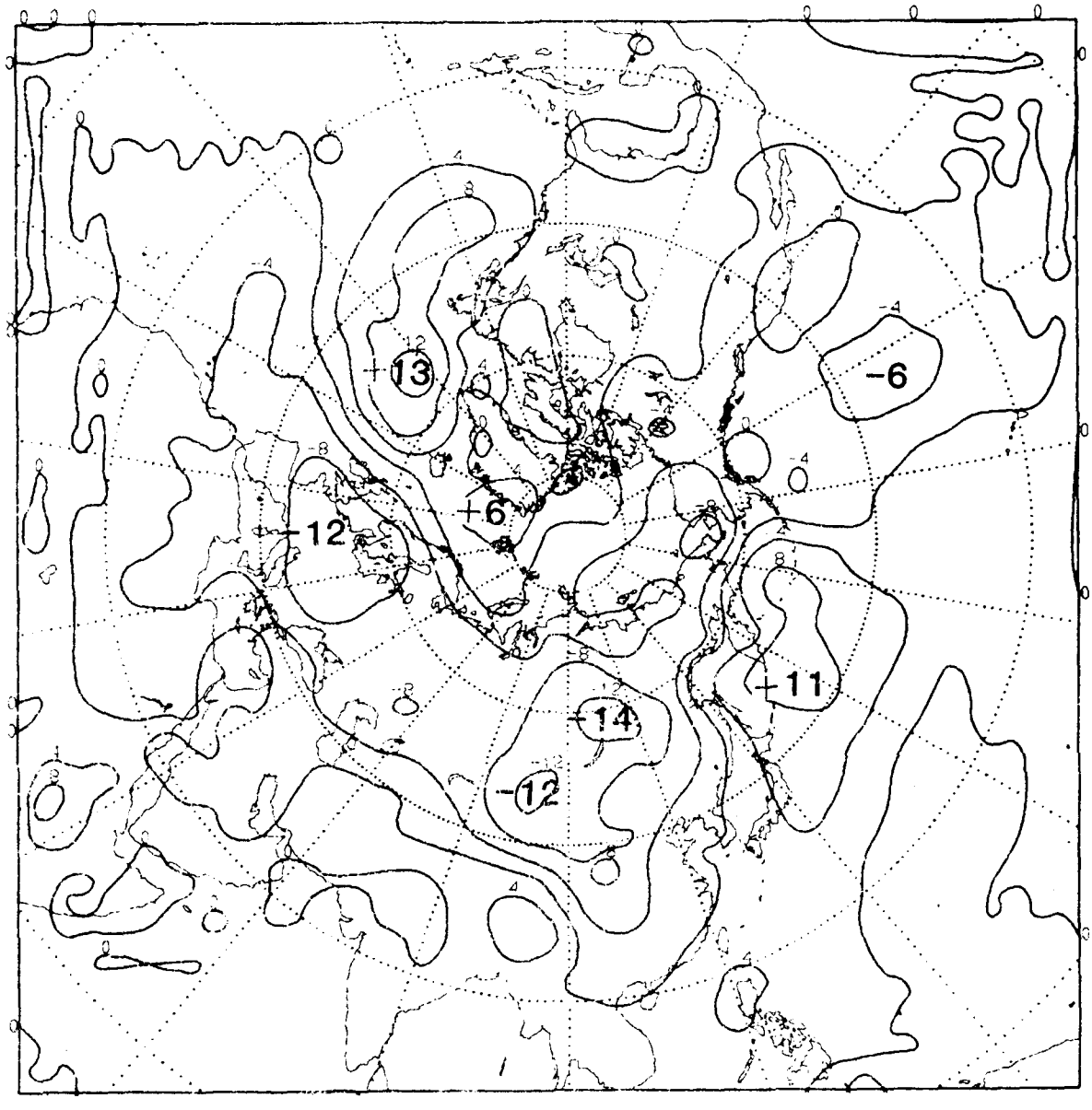

Fig. 15b. Systematic error (mb) of the SAT 48 hour forecast of sea level pressure for 16-28 December 1987 at 1200 GMT.

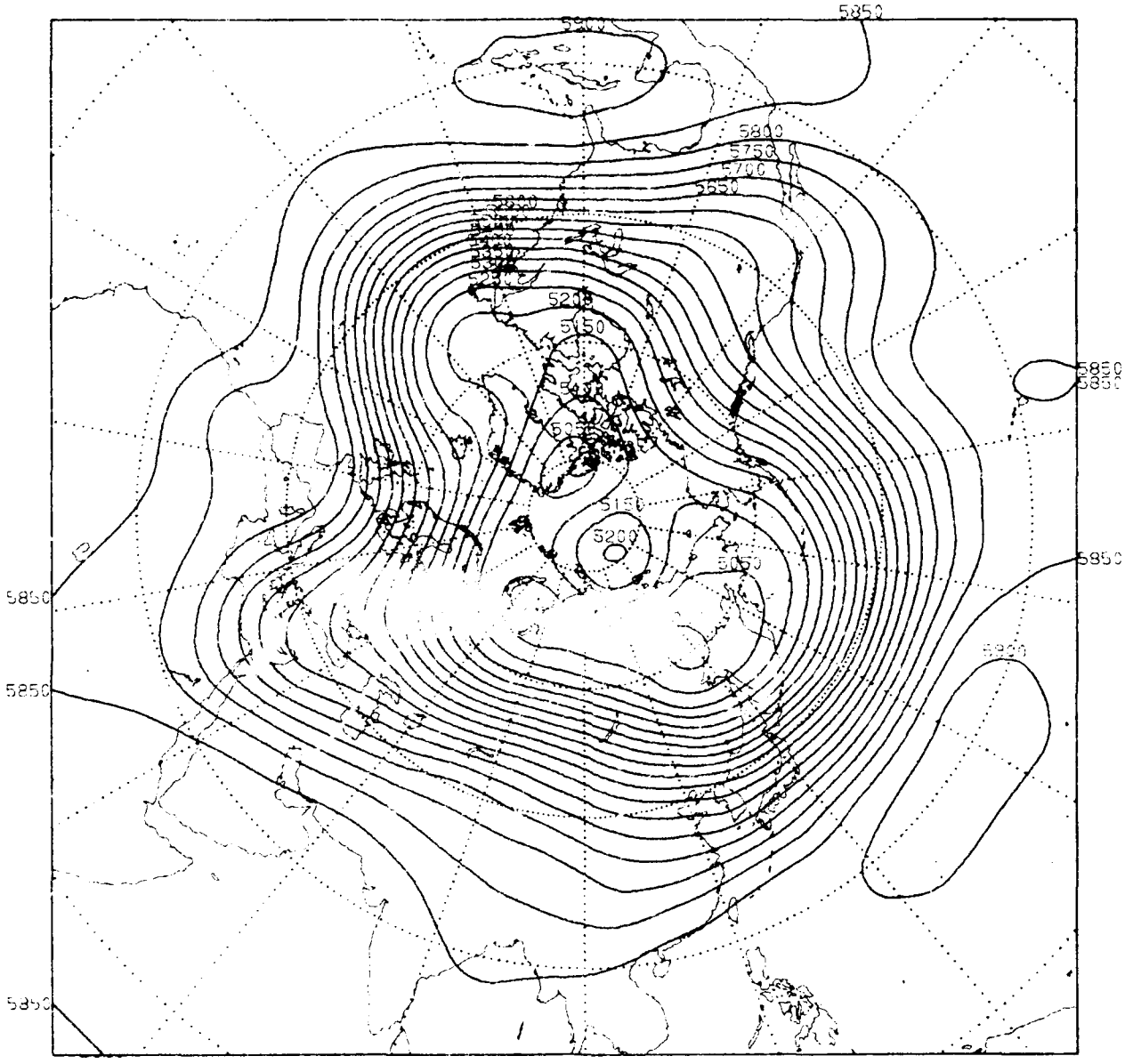

Fig. 16a. Average NOSAT 500 mb height analysis (m) for 16-28 December 1987 at 1200 GMT.

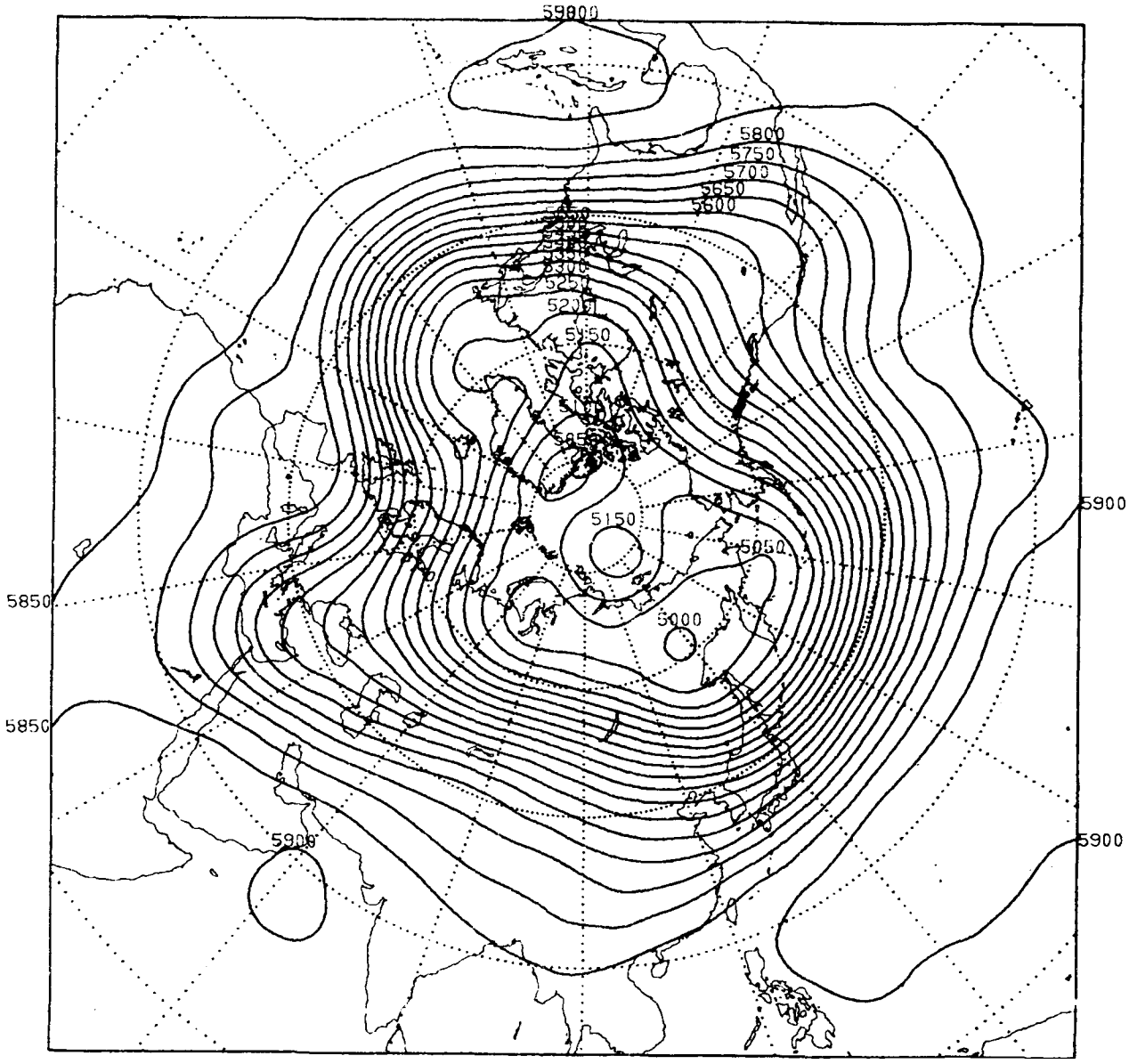

Fig. 16b. Average SAT 500 mb height analysis (m) for 16-28 December 1987 at 1200 GMT.

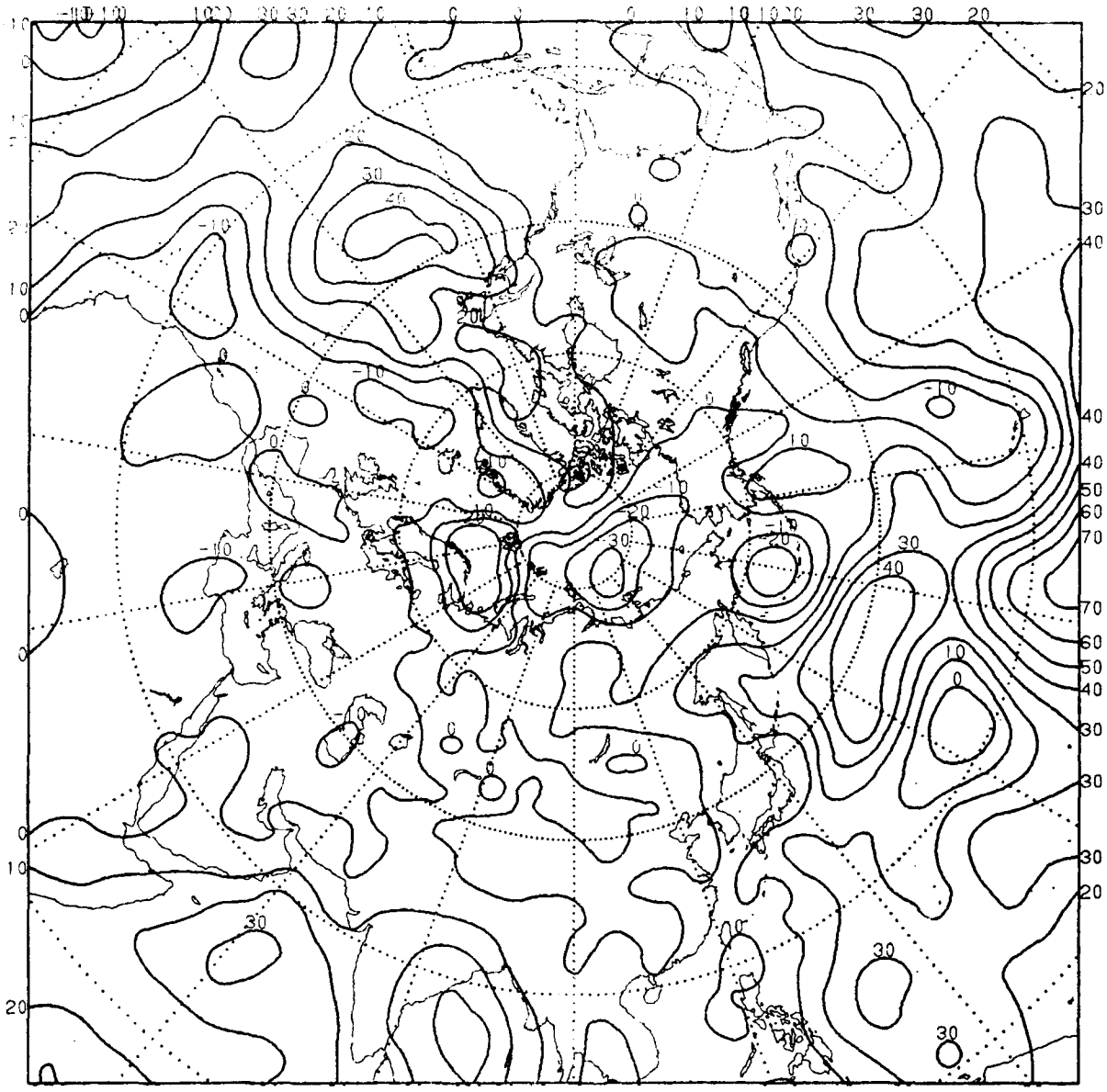

Fig. 17. Differences between SAT and NOSAT (SAT - NOSAT) average 500 mb analyses (m) for 16-28 December 1987 at 1200 GMT.

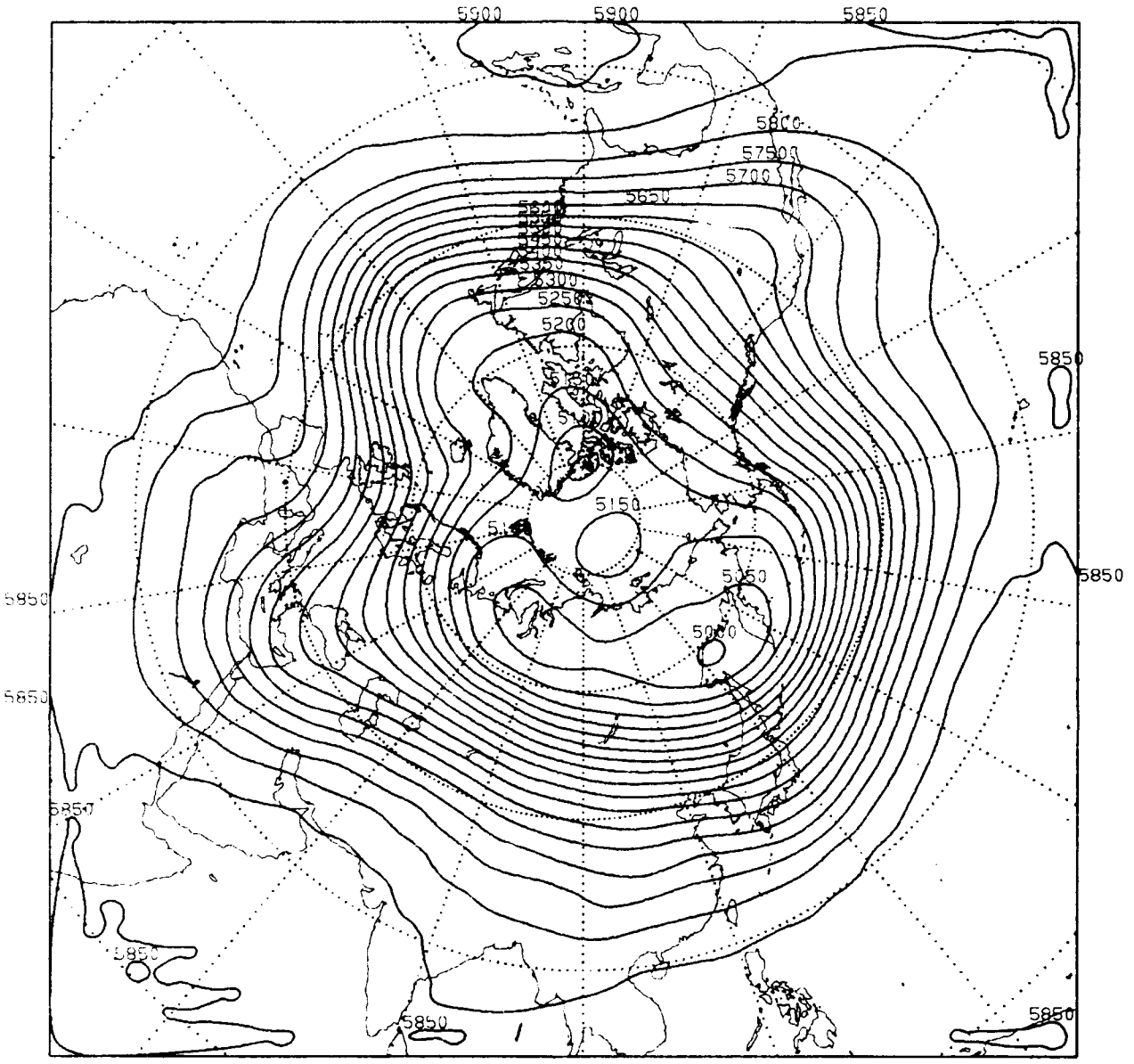

Fig. 18a. Average 48 hour NOSAT forecast of 500 mb height (m) for 16-28 December 1987 at 1200 GMT.

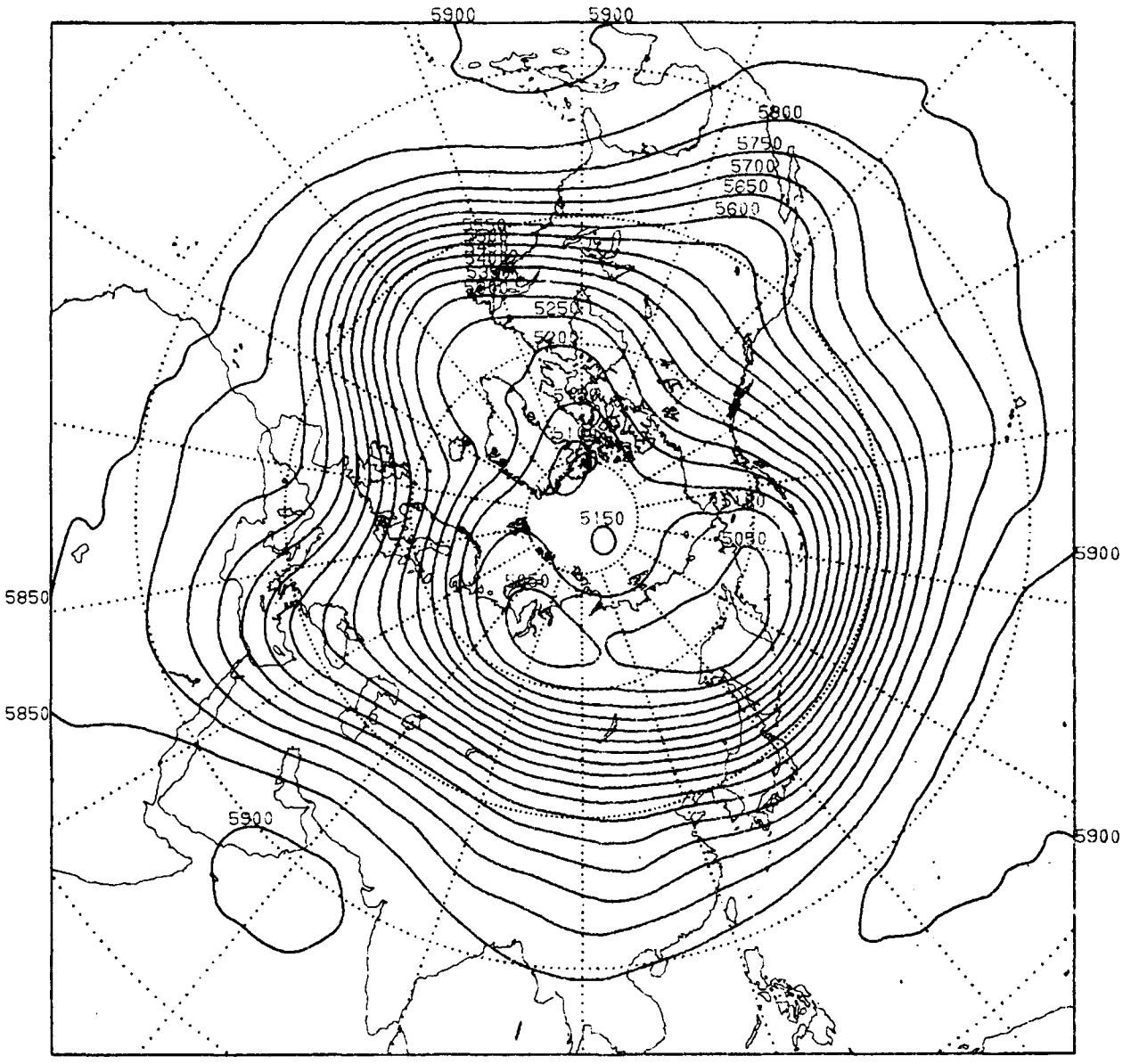

Fig. 18b. Average 48 hour SAT forecast of 500 mb height (m) for 16-28 December 1987 at 1200 GMT.

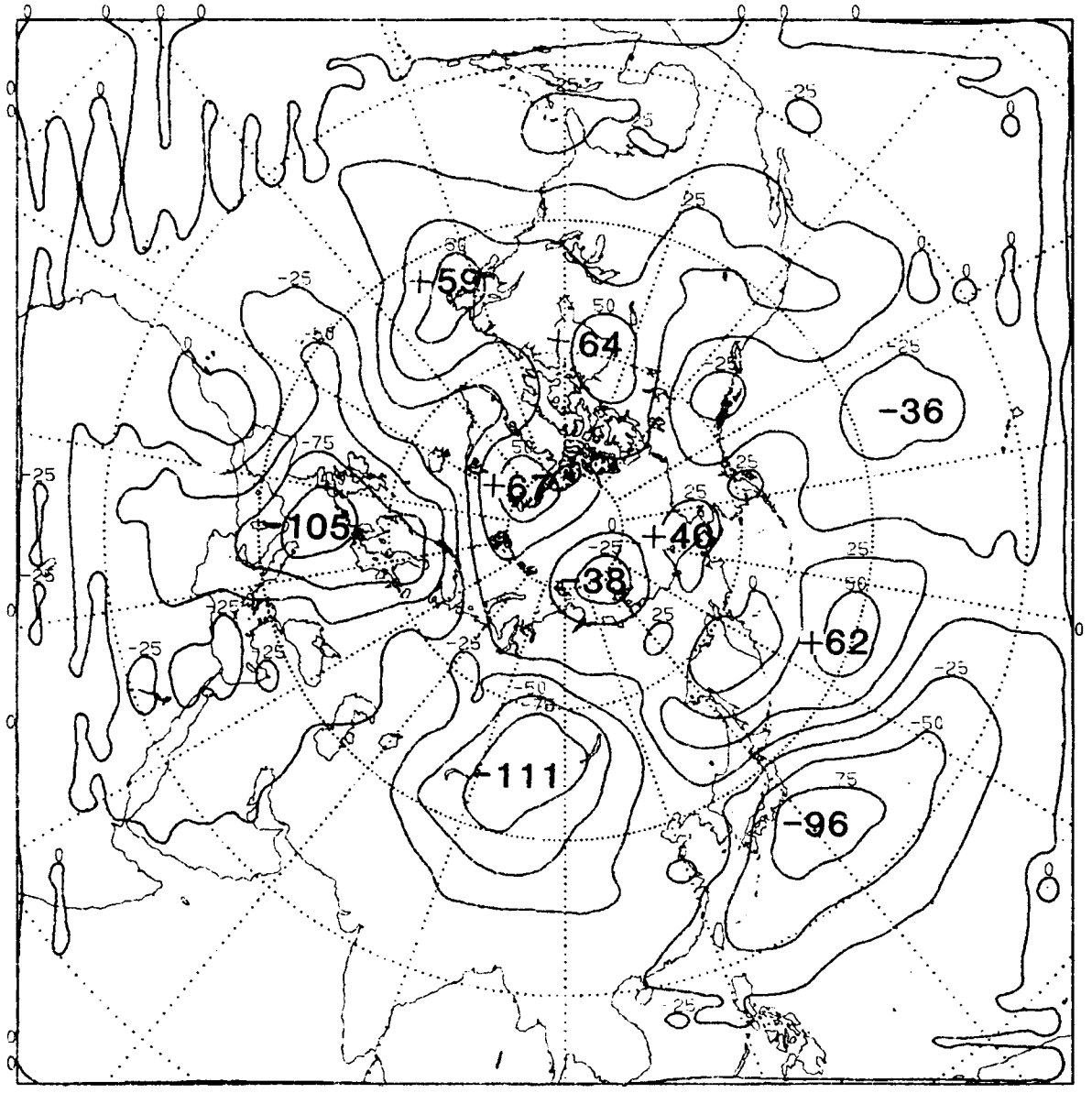

Fig. 19a. Systematic error (m) of the NOSAT 48 hour forecast of 500 mb height for 1[illegible]-28 December 1987 at 1200 GMT.

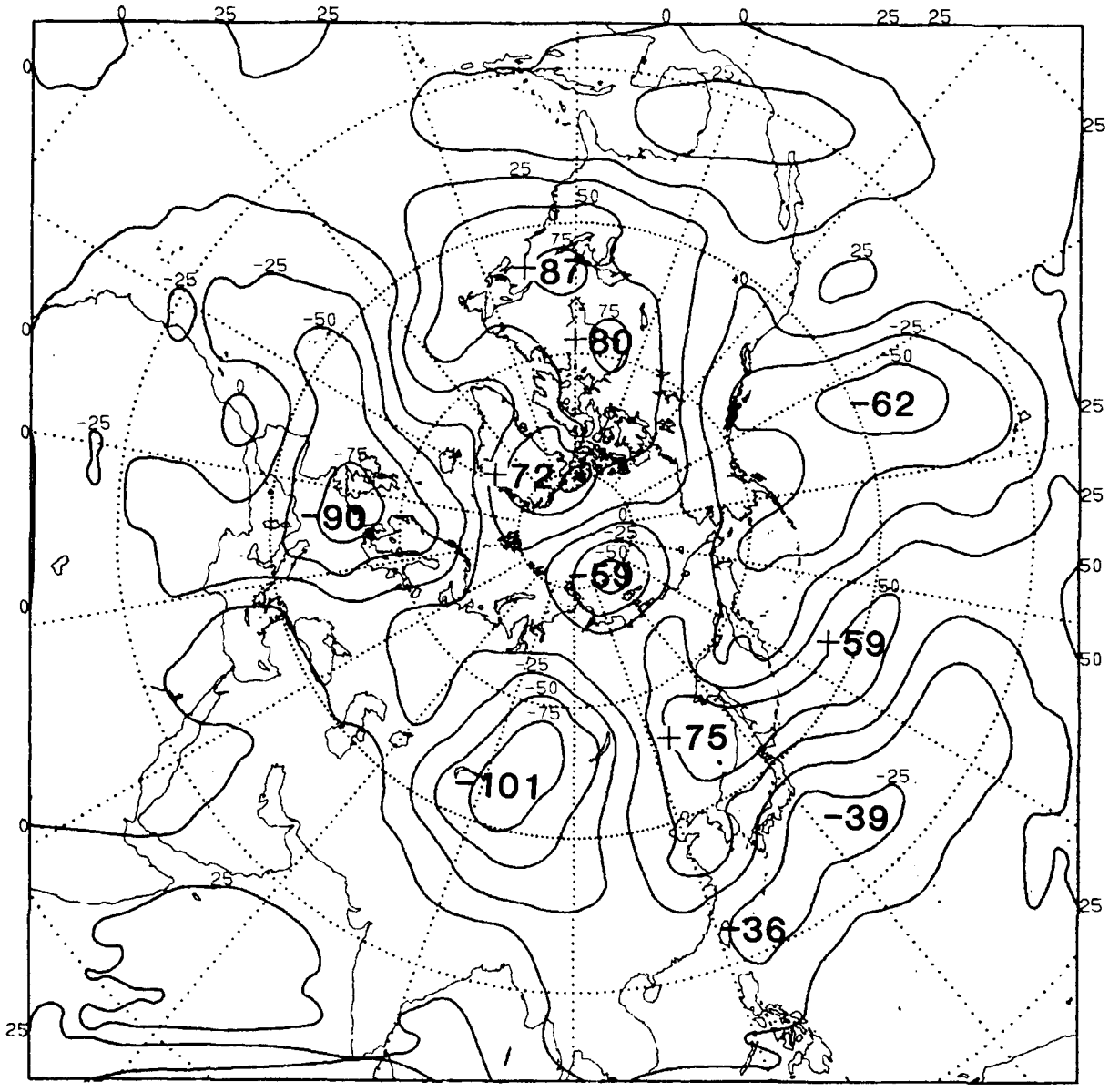

Fig. 19b. Systematic error (m) of the SAT 48 hour forecast of 500 mb height for 16-28 December 1987 at 1200 GMT.

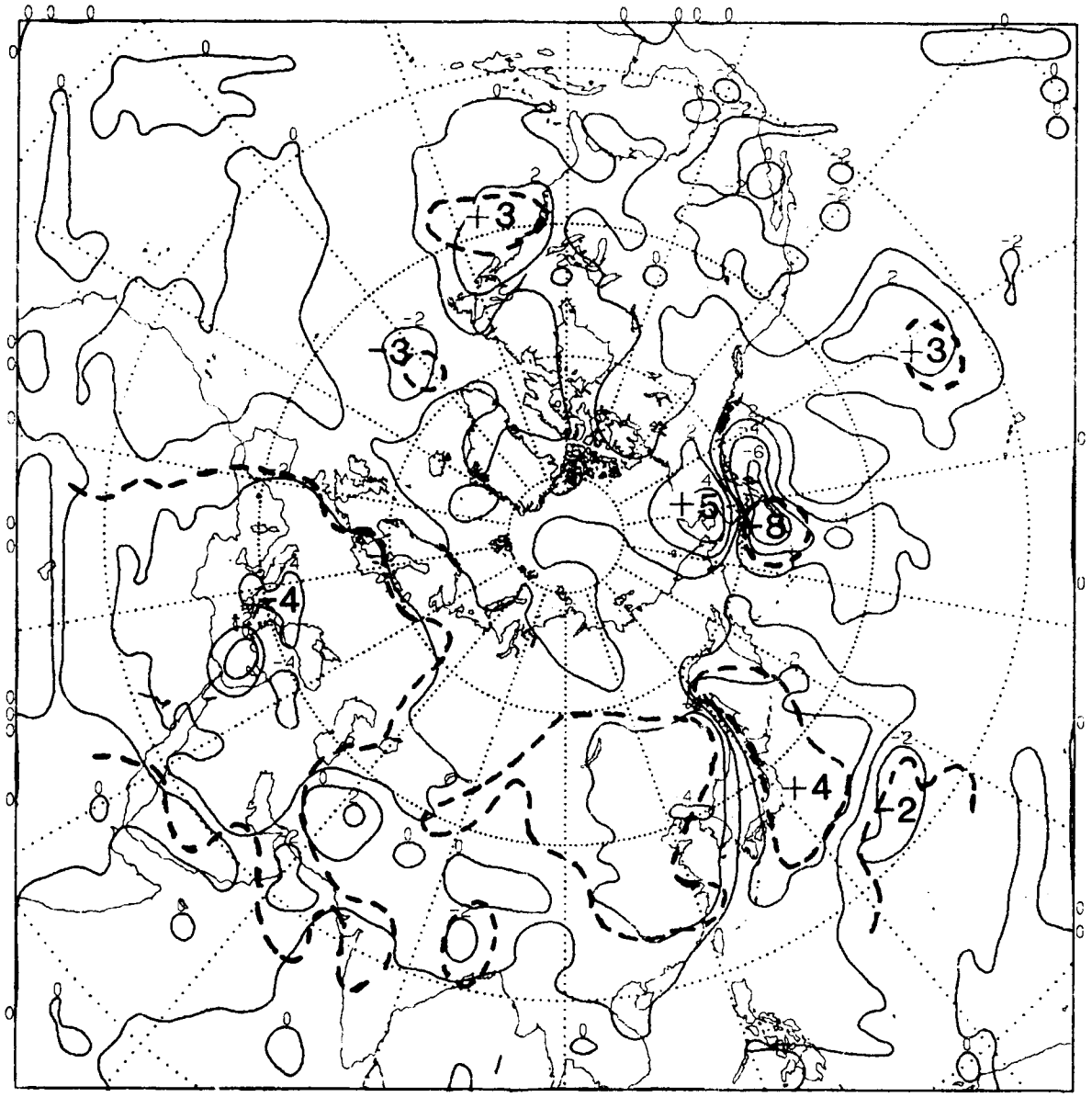

Fig. 20a. Differences between absolute values of SAT and NOSAT systematic errors (|SAT| - |NOSAT|) of 48 hour forecasts of sea level pressure for 16-28 December 1987 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

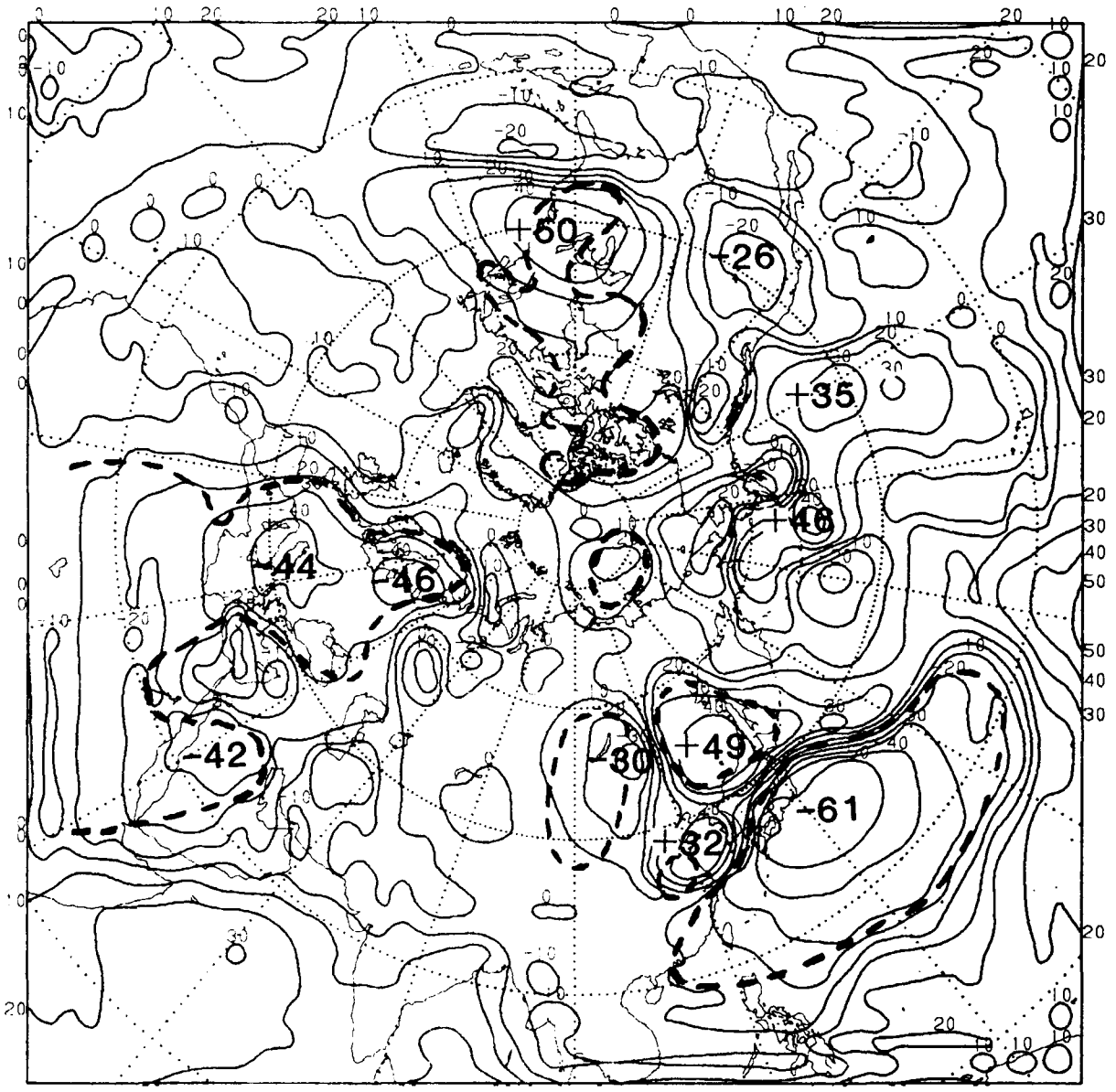

Fig. 20b. Differences between absolute values of SAT and NOSAT systematic errors (|SAT| - |NOSAT|) of 48 hour forecasts of 500 mb height (m) for 16-28 December 1987 at 1200 GMT. Heavy dashed lines enclose statistically significant values.

☆ U.S. GOVERNMENT PRINTING OFFICE 1990—261-915 20605

(Continued from inside front cover)

NESDIS 19 Earth Observations and the Polar Platform. John H. McElroy and Stanley R. Schneider, January 1985. (PB85 177624/AS)

NESDIS 19 The Space Station Polar Platform: Intergrating Research and Operational Missions. John H. McElroy and Stanley R. Schneider, January 1985. (PB85 195279/AS)

NESDIS 20 An Atlas of High Altitude Aircraft Measured Radiance of White Sands, New Mexico, in the 450-1050nm Band. Gilbert R. Smith, Robert H. Levin and John S. Knoll, April 1985. (PB85 204501/AS)

NESDIS 21 High Altitude Measured Radiance of White Sands, New Mexico, in the 400-2000nm Band Using a Filter Wedge Spectrometer. Gilbert R. Smith and Robert H. Levin, April 1985. (PB85 206084/AS)

NESDIS 22 The Space Station Polar Platform: NOAA Systems Considerations and Requirements. John H. McElroy and Stanley R. Schneider, June 1985. (PB86 6109246/AS)

NESDIS 23 The Use of TOMS Data in Evaluating and Improving the Total Ozone from TOVS Measurements. James H. Lienesch and Prabhat K.P. Pandey, July 1985. (PB86 108412/AS)

NESDIS 24 Satellite-Derived Moisture Profiles. Andrew Timchalk, April 1986. (PB86 232923/AS)

NESDIS 25 reserved

NESDIS 26 Monthly and Seasonal Mean Outgoing Longwave Radiation and Anomalies. Arnold Gruber, Marylin Varnadore, Phillip A. Arkin, and Jay S. Winston, October 1987. (PB87160545/AS)

NESDIS 27 Estimation of Broadband Planetary Albedo from Operational Narrowband Satellite Measurements. James Wydick, April 1987. (PB88-107644/AS)

NESDIS 28 The AVHRR/HIRS Operational Method for Satellite Based Sea Surface Temperature Determination. Charles Walton, March 1987. (PB88-107594/AS)

NESDIS 29 The Complementary Roles of Microwave and Infrared Instruments in Atmospheric Sounding. Larry McMillin, February 1987. (PB87 134917/AS)

NESDIS 30 Planning for Future Generational Sensors and Other Priorities. James C. Fischer, June 1987. (PB87 220802/AS)

NESDIS 31 Data Processing Algorithms for Inferring Stratospheric Gas Concentrations from Balloon-Based Solar Occultation Data. I-Lok Chang (American University) and Michael P. Weinreb, April 1987. (PB87 196424)

NESDIS 32 Precipitation Detection with Satellite Microwave Data. Yang Chenggang and Andrew Timchalk, June 1988. (PB88-240239)

NESDIS 33 An Introduction to the GOES I-M Imager and Sounder Instruments and the GVAR Retransmission Format. Raymond J. Komajda (Mitre Corp) and Keith McKenzie, October 1987. (PB88-132709)

NESDIS 34 Balloon-Based Infrared Solar Occultation Measurements of Stratospheric O3, H2O, HNo3, and CF2C12. Michael P. Weinreb and I-Lok Chang (American University), September 1987. (PB88-132725)

NESDIS 35 Passive Microwave Observing From Environmental Satellites, A Status Report Based on NOAA's June 1-4, 1987, Conference in Williamsburg, Virginia. James C. Fischer, November 1987. (PB88-208236)

NESDIS 36 Pre-Launch Calibration of Channels 1 and 2 of the Advanced Very High Resolution Radiometer. C.R. Nagaraja Rao, October 1987. (PB88-157169 A/S)

NESDIS 39 General Determination of Earth Surface Type and Cloud Amount Using Multispectral AVHRR Data. Irwin Ruff and Arnold Gruber, February 1988. (PB88-199195/AS)

NESDIS 40 The GOES I-M System Functional Description. Carolyn Bradley (Mitre Corp), November 1988.

NESDIS 41 Report of the Earth Radiation Budget Requirements Review - 1987 Rosslyn, Virginia, 30 March - 3 April 1987. L.L. Stowe (Editor), June 1988.

NESDIS 42 Simulation Studies of Improved Sounding Systems. H. Yates, D. Wark, H. Aumann, N. Evans, N. Phillips, J. Sussking, L. McMillin, A. Goldman, M. Chahine, and L. Crone, February 1989.

NESDIS 43 Adjustment of Microwave Spectral Radiances of the Earth to a Fixed Angle of Progagation. D. Q. Wark, December 1988. (PB89-162556/AS)

NESDIS 44 Educator's Guide for Building and Operating Environmental Satellite Receiving Stations. R. Joe Summers, Chambersburg Senior High, February 1989.

NESDIS 45 Final Report on the Modulation and EMC Considerations for the HRPT Transmission System in the Post NOAA-M Polar Orbiting Satellite ERA. James C. Fischer (Editor), June 1989. (PB89-223812/AS)

NESDIS 46 MECCA Program Documentation. Kurt W. Hess, September 1989.

NOAA SCIENTIFIC AND TECHNICAL PUBLICATIONS

The National Oceanic and Atmospheric Administration was established as part of the Department of Commerce on October 3, 1970. The mission responsibilities of NOAA are to assess the socioeconomic impact of natural and technological changes in the environment and to monitor and predict the state of the solid Earth, the oceans and their living resources, the atmosphere, and the space environment of the Earth.

The major components of NOAA regularly produce various types of scientific and technical information in the following kinds of publications:

PROFESSIONAL PAPERS—Important definitive research results, major techniques, and special investigations.

CONTRACT AND GRANT REPORTS—Reports prepared by contractors or grantees under NOAA sponsorship.

ATLAS—Presentation of analyzed data generally in the form of maps showing distribution of rainfall, chemical and physical conditions of oceans and atmosphere, distribution of fishes and marine mammals, ionospheric conditions, etc.

TECHNICAL SERVICE PUBLICATIONS—Reports containing data, observations, instructions, etc. A partial listing includes data serials; prediction and outlook periodicals; technical manuals, training papers, planning reports, and information serials; and miscellaneous technical publications.

TECHNICAL REPORTS—Journal quality with extensive details, mathematical developments, or data listings.

TECHNICAL MEMORANDUMS—Reports of preliminary, partial, or negative research or technology results, interim instructions, and the like.

U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL ENVIRONMENTAL SATELLITE, DATA, AND INFORMATION SERVICE
Washington, D.C. 20233